

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 00:01:12 ; Search time 819 Seconds

(without alignments)
10446.729 Million cell updates/sec

Title: US-09-980-364-1

Perfect score: 2014

Sequence: 1 gtctactctctcttcttaag.....gttgataatttcagact 2014

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2014	100.0	2014	4 AAC91398	AAC91398 Brassica
2	1867	92.7	2011	4 AAC91399	AAC91399 Brassica
3	726	36.0	4873	4 AAC91400	AAC91400 Brassica
4	420.6	20.9	2463	7 ACC00803	ACC00803 Glycine m
5	413.6	20.5	1824	7 ACC00802	ACC00802 Glycine m
6	413.6	20.5	2598	7 ACC00801	ACC00801 Glycine m
7	399.2	19.8	1749	4 AAC93248	AAC93248 Cellulose
8	385.8	19.2	1990	7 ACC00806	ACC00806 Glycine m
9	377.4	18.7	1818	9 ADD30878	ADD30878 Plant yie
10	377.4	18.7	1818	9 ADD30822	ADD30822 Plant yie
11	377.4	18.7	1818	9 ADD30333	ADD30333 Plant yie
12	357.4	17.7	1935	7 ACC00848	ACC00848 Gossypium
13	352	17.5	2168	7 ACC00804	ACC00804 Glycine m
14	347.8	17.3	1290	7 ACC00805	ACC00805 Glycine m
15	347.8	17.3	1954	7 ACC00850	ACC00850 Glycine m
16	335.2	16.6	1878	7 ACC00846	ACC00846 Glycine m
17	334	16.6	1874	7 ACC00799	ACC00799 Glycine m
18	332.2	16.5	2052	9 ADD30980	ADD30980 Plant yie
19	330.6	16.4	1668	3 AAC43005	AAC43005 Arabidops
20	330.6	16.4	2148	3 AAA59220	AAA59220 cDNA enco
21	330.6	16.4	2148	7 ABX13437	ABX13437 A. thalia
22	330.2	16.4	1959	7 ACC00847	ACC00847 Catalpa s
23	327.6	16.3	2374	7 ACC00849	ACC00849 Zea mays

ALIGNMENTS

RESULT 1

AAC91398
ID AAC91398 standard; cDNA; 2014 BP.

XX
AC AAC91398;

DT 19-MAR-2001 (first entry)

XX
DE Brassica napus BNM3A cDNA.

XX
KW Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3A;

XX
KW asexually derived embryo production; increased regenerative capacity; ss.

XX
OS Brassica napus.

XX
PN EP1057891-A1.

XX
PD 06-DEC-2000.

XX
PF 02-JUN-1999; 99EP-00201745.

XX
PR 02-JUN-1999; 99EP-00201745.

XX
PA (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
(MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

XX
PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
Boutillier K;

XX
WPI: 2001-026720/04.

XX
P-PSDB; AAB50872.

XX
Brassica napus BNM3 DNA sequences, useful for transforming host cells to produce asexually derived embryos.

XX
Claim 4; Fig 2; 51pp; English.

XX
CC The present sequence is given in a specification relating to a Brassica napus microspore embryo (BNM) gene obtained during the induction of a microspore embryogenesis. The isolated DNA molecule can be used in a vector that is used to transform plant cells. The vector can be used in methods to produce asexually derived embryos, for modifying the regenerative capacity of a plant, and for producing an apomictic plant

XX
SQ Sequence 2014 BP; 596 A; 411 C; 462 G; 545 T; 0 U; 0 Other;

end

XX	19-MAR-2001 (first entry)	QY	490	GAGGAGGACGGTGTGGTGGCTCACTAGGACTTTTCGATGATAAAGACATGCTGTGAA	549
DT		Db	421	GAGGAGGACGGTGTGGTGGCTCACTAGGACTTTTCGATGATAAAGACATGCTGTGAA	480
XX	Brassica napus BNM3B CDNA.	QY	550	ATCAACCCGTGGATAATGTTGATAATCAAGAAATGCAATGCTGCAAAAGGCTGTCCC	609
XX	Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3B;	Db	481	ATCAACCCGTGGATAATGTTGATAATCAAGAAATGCAATGCTGCAAAAGGCTGTCCC	540
KW	asexually derived embryo production; increased regenerative capacity; ss.	QY	610	TCTCAATGAACACTCTACTCTTCTTGTGATAACCAACGACGACGCAATACCAACGTTTG	669
XX	Brassica napus.	Db	541	TCTCAATGAACACTCTACTCTTCTTGTGATAACCAACACTACAGCAGTAACAACCTTGTG	600
XX	BP1057891-A1.	QY	670	CCCAAGGGAAGACTATTGATAGCTTGAAGCTACACCGAAGAAACTATTTCAGAGATT	729
XX	06-DEC-2000.	Db	601	CCCAAGGGAAGACTATTGATGATAGCTTGAAGCTACACCGAAGAAACTATTTCAGAGATT	660
XX	02-JUN-1999; 99BP-00201745.	QY	730	TTGGACAGAGCAGCTCTATATACCCCGGTGTTACAAGGCATTCGGTGACAGGAAGATATG	789
XX	02-JUN-1999; 99BP-00201745.	Db	661	TTGGACAGAGCAGCTCTATATACCCCGGTGTTACAAGGCATTCGGTGACAGGAAGATATG	720
XX	(CPRO-) CPRO-DLO CENT PLANTENVERDEDELINGS REPROD.	QY	790	AGGCACATTTATGGGATAATAGTTGTAAGAGAGAGGCCAAACGCGCAAGGAAGACAAG	849
PA	(MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.	Db	721	AGGCACATTTATGGGATAATAGTTGTAAGAGAGAGGCCAAACGCGCAAGGAAGACAAG	780
XX	Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;	QY	850	TTTATTTGGGAGGTTATGACAAAGAAAGAGAGAGGCTAGGGCTTATGATTTAGCCGAC	909
PI	Boutillier K;	Db	781	TTTATTTGGGAGGTTATGACAAAGAAAGAGAGGCTAGGGCTTATGATTTAGCCGAC	840
XX	WPI; 2001-026720/04.	QY	910	TCAAGTATTTGGGAAACCCACTACTACTACTACTACTACTACTACTACTACTACTACTACT	969
DR	P-PSDB; AAB50873.	Db	841	TCAAGTATTTGGGAAACCCACTACTACTACTACTACTACTACTACTACTACTACTACTACT	900
XX	Brassica napus BNM3 DNA sequences, useful for transforming host cells to	QY	970	TAGAGAGATCAAGCACATGACAGGCAAGAGTATGTTGCTCTCACTGCGCAGGAAAGTA	1029
PT	produce asexually derived embryos.	Db	901	TAGAGAGATCAAGCACATGACAGGCAAGAGTATGTTGCTCTCACTGCGCAGGAAAGTA	960
XX	Claim 4; Fig 2; Sipp; English.	QY	1030	FTGGTTTCTCTGTTGTCATCGATTATCTGTGAGTAAACAGACATCACCACATGAA	1089
CC	The present sequence is given in a specification relating to a Brassica	Db	961	FTGGTTTCTCTGTTGTCATCGATTATCTGTGAGTAAACAGACATCACCACATGAA	1020
CC	napus microspore embryo (BMN) gene obtained during the induction of	QY	1090	GATGCGCAAGCTAGGATAGGAAGAGTGCCTGCTTACAAAGACCTCTACTTGGGAACCTTTG	1149
CC	microspore embryogenesis. The isolated DNA molecule can be used in a	Db	1021	GATGCGCAAGCTAGGATAGGAAGAGTGCCTGCTTACAAAGACCTCTACTTGGGAACCTTTG	1080
CC	vector that is used to transform plant cells. The vector can be used in	QY	1150	GCACAACAAGAAAGCTCCAGAGGATACGACATTCGCGGCATCAAAATTCAGAGATTAA	1209
CC	methods to produce asexually derived embryos, for modifying the	Db	1081	GCACAACAAGAAAGCTCCAGAGGATACGACATTCGCGGCATCAAAATTCAGAGATTAA	1140
CC	regenerative capacity of a plant, and for producing an apomictic plant	QY	1210	CGGAGTGAACCTTCGACATGACATCAACGTTTAAAGCATCTCTGAAAGCCCTA	1269
XX	Sequence 2011 BP; 610 A; 395 C; 462 G; 544 T; 0 U; 0 Other;	Db	1141	CGGAGTGAACCTTCGACATGACATCAACGTTTAAAGCATCTCTGAAAGCCCTA	1200
QY	Query Match 92.7%; Score 1867; DB 4; Length 2011;	QY	1270	GTCTTCTCTATTGTTAGCGCCGCAAAACGCTCTCAAGGAGGCTAAACGCTCCGTTCCAAAGTA	1329
Db	Best Local Similarity 97.8%; Pred. No. 0;	Db	1201	GTCTTCTCTATTGTTAGCGCCGCAAAACGCTCTCAAGGAGGCTAAACGCTCCGTTCCAAAGTA	1260
Db	Matches 1903; Conservative 0; Mismatches 40; Indels 2; Gaps 1;	QY	1330	TGATGATGATCAGTAACTTAACGTTTCAGAGTGAACATAGTCTAGCGGTTGGCAAAAGC	1389
QY	70 TTCTTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAACCTCAATGAATAA	Db	1261	TGATGATGATCAGTAACTTAACGTTTCAGAGTGAACATAGTCTAGCGGTTGGCAAAAGC	1320
Db	1 TTCTTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTTTCTGAACTCGATGAATAA	QY	1390	CTGCGGTTTCAGGATCATCAGGAGTAGATTTCAGGCTTATTCACCAACATCAAGAGAGGT	1449
QY	130 ACTGTTAGGCTTTTCTCTCTCTCTTTATGAACAAATCAACATCGTAAGGAGCTTACT	Db	1321	CTGCGGTTTCAGGATCATCAGGAGTAGATTTCAGGCTTATTCACCAACATCAAGAGAGGT	1380
Db	61 ACTGTTAGGCTTTTCTCTCTCTCTCTTTATGAACAAATCAACATCGTAAGGAGCTTCT	QY	1450	ACAATGGTTATTATTAACAATGGAGAACTTGTCTTCGAGAGTCTAGGCTTGTTC	1509
QY	190 CTTCCACACACAAACCGTGTAGATGTCGCGGAGAGTACTGTAGCATCCGACCGCTG	Db	1381	ACAATGGTTATTATTAACAATGGAGAACTTGTCTTCGAGAGTCTAGGCTTGTTC	1440
Db	121 CTTCCACACACAAACCGTGTAGATGTCGCGGAGAGTACTGTAGCATCCGACCGCTG	QY	1510	ACAATGGTTATTATTAACAATGGAGAACTTGTCTTCGAGAGTCTAGGCTTGTTC	1569
QY	250 CTTCCAGTGTCTTACGACCAATCAAAACATCGTTTCTCTCTCTCTCTCTCTCTCTCT	Db	1441	ACAATGGTTATTATTAACAATGGAGAACTTGTCTTCGAGAGTCTAGGCTTGTTC	1500
Db	181 CTTCCAGTGTCTTACGACCAATCAAAACATCGTTTCTCTCTCTCTCTCTCTCTCTCT	QY	1570	ATCATCAAAAGTCTCTGTTTCGGATGATTCGGTTACTGTTTGTGGAAATGTTGTGTTATG	1629
QY	310 ATGCTTTTACAGAGACAAATAGTCACTCCCGAGATTGGACATCAATGGTTGTCAT				
Db	241 ATGCTTTTACAGAGACAAATAGTCACTCCCGAGATTGGACATCAATGGTTGTCAT				
QY	370 GCAATAACATCCAAACGATGAGCAAGATGACCAAGCTTGAAGTTTCTCTGCGCGCA				
Db	301 GTAATAACATCCAAACGATGAGCAAGATGACCAAGCTTGAAGTTTCTCTGCGCGCA				
QY	430 CCACCAAGATTACAAACCAACCAACGTTGAGATGGAAGTGGAGTGGCTGTTATG				
Db	361 CCACCAAGATTACAAACCAACCAACGTTGAGATGGAAGTGGAGTGGCTGTTATG				

Db 1501 ATCATCAAGATTCTGTTTCAGATGATTCGGTACTGTTGTGGAAATGTTGTTGTTATG 1560
 QY 1630 GTGGTTATCAAGATTTCGAGCCCGGTTAACTGCGATGCTACGCTGCTAGTCAAGTTG 1689
 Db 1561 GTGGTTATCAAGATTTCGAGCCCGGTTAACTGCGATGCTACGCTGCTAGTCAAGTTG 1620
 QY 1690 ATTATAACGCAAGAAACCAATTATTAATTTGCTCAGCAGCAGCAGCAGCAGCAGCAG 1749
 Db 1621 ACTATAACGCAAGAAACCAATTATTAATTTGCTCAGCAGCAGCAGCAGCAGCAGCAG 1680
 QY 1750 GTGGAGATTTCCTCCCGCGCAATGACGAATAATGTTGGCTCTAATATGTTATACCATGGG 1809
 Db 1681 GAGGAGATTTCCTCCCGCGCAATGACGAATAATGTTGGCTCTAATATGTTATACCATGGG 1740
 QY 1810 AAGGTGGTGGAGAACTGCTCCAACTTACAGTTTGGAAACGACAAATPAGAAAAATAGT 1869
 Db 1741 AAGGTGGTGGAGAACTGCTCCAACTTACAGTTTGGAAACGACAAATPAGAAAAATAGT 1800
 QY 1870 TAAAGATCTTTAGTTATATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1929
 Db 1801 TAAAGATCTTTAGTTATATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
 QY 1930 GTTTTCTTCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1989
 Db 1861 GTTTTCTTCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1918
 QY 1990 CATTAGTTGGATAAATTTTCAGACT 2014
 Db 1919 CATTAGTTGGATAAATTTTCGACT 1943

RESULT 3

AAC91400
 ID AAC91400 standard; DNA; 4873 BP.
 XX AC AAC91400;

19-MAR-2001 (first entry)

Brassica napus BNM3A DNA sequence.

Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3A;
 asexually derived embryo production; increased regenerative capacity; ds.

Brassica napus.

EP1057891-A1.

06-DEC-2000.

02-JUN-1999; 99EP-00201745.

02-JUN-1999; 99EP-00201745.

(CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
 (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
 Boutillier K;

WPI; 2001-026720/04.
 P-PSDB; AAB50872.

Brassica napus BNM3 DNA sequences, useful for transforming host cells to
 produce asexually derived embryos.

Claim 1; Page 32-33; 51pp; English.

The present sequence is given in a specification relating to a Brassica
 napus microspore embryo (BNM) gene obtained during the induction of
 microspore embryogenesis. The isolated DNA molecule can be used in a
 vector that is used to transform plant cells. The vector can be used in

CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant
 XX Sequence 4873 BP; 1596 A; 856 C; 907 G; 1514 T; 0 U; 0 Other;
 Query Match 36.0%; Score 726; DB 4; Length 4873;
 Best Local Similarity 100.0%; Pred. No. 1.4e-175;
 Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1149 GGCACACAAAGAGAGCTGCAGAGGATACGACATTCGGCCATCAAAATTCAGAGGATTA 1208
 Db 4148 GGCACACAAAGAGAGCTGCAGAGGATACGACATTCGGCCATCAAAATTCAGAGGATTA 4207
 QY 1209 ACCGAGTGACTAACTTCGACATGAACAGATACAACTGTTAAAGCAATCTCTGAAAGCCCT 1268
 Db 4208 ACCGAGTGACTAACTTCGACATGAACAGATACAACTGTTAAAGCAATCTCTGAAAGCCCT 4267
 QY 1269 AGTCTTCCTATTGGTAGCGCCAAAACGTTCTCAAGGAGGCTAACCGTCCGGTTCCTCAAGT 1328
 Db 4268 AGTCTTCCTATTGGTAGCGCCAAAACGTTCTCAAGGAGGCTAACCGTCCGGTTCCTCAAGT 4327
 QY 1329 ATGATGATGATCAGTAATAAAGCTTTTCAGAGAGTGAAGATAGTGTGCTAGCGTTGCAAAAC 1388
 Db 4328 ATGATGATGATCAGTAATAAAGCTTTTCAGAGAGTGAAGATAGTGTGCTAGCGTTGCAAAAC 4387
 QY 1389 GTCGGGTTTCAGATCATCAGGAGTGAATTTGAGCTTATTCACCAACATCAGAGAGG 1448
 Db 4388 GTCGGGTTTCAGATCATCAGGAGTGAATTTGAGCTTATTCACCAACATCAGAGAGG 4447
 QY 1449 TACAATGTTTATTATTAATGGAGAACTTGTCTTCGGAGAGTGTAGGCTTGTGTTTC 1508
 Db 4448 TACAATGTTTATTATTAATGGAGAACTTGTCTTCGGAGAGTGTAGGCTTGTGTTTC 4507
 QY 1509 AAACAAGAGGATGATCAACACCATTTCTTGAGCAACACGAGAGCTCATGACTAATATC 1568
 Db 4508 AAACAAGAGGATGATCAACACCATTTCTTGAGCAACACGAGAGCTCATGACTAATATC 4567
 QY 1569 GATCATCAAAAGTTCTGTTTCGGATGATTTCGGTTACTGTTTGGAAATGTTGTTGGTTAT 1628
 Db 4568 GATCATCAAAAGTTCTGTTTCGGATGATTTCGGTTACTGTTTGGAAATGTTGTTGGTTAT 4627
 QY 1629 GGTGTTATCAAGGATTTGACGCCCGGTTAACTGCGATGCTAGCTGCTAGTGAAGTTT 1688
 Db 4628 GGTGTTATCAAGGATTTGACGCCCGGTTAACTGCGATGCTAGCTGCTAGTGAAGTTT 4687
 QY 1689 GATTATAACGCAAGAAACCAATTATTACTTGTCTAGCAGCAGCAGCAGCAGCAGCAGCAG 1748
 Db 4688 GATTATAACGCAAGAAACCAATTATTACTTGTCTAGCAGCAGCAGCAGCAGCAGCAGCAG 4747
 QY 1749 GGTGAGATTTTCCCGCGCAATGACGAATAATGTTGGCTCTAATATGTTATTCATGGG 1808
 Db 4748 GGTGAGATTTTCCCGCGCAATGACGAATAATGTTGGCTCTAATATGTTATTCATGGG 4807
 QY 1809 GAAGGTGGTGGAGAGTTGCTCCAACTTACAGTTTGGAAACGACAAATPAGAAAAATAG 1868
 Db 4808 GAAGGTGGTGGAGAGTTGCTCCAACTTACAGTTTGGAAACGACAAATPAGAAAAATAG 4867
 QY 1869 TTAAG 1874
 Db 4868 TTAAG 4873

RESULT 4

ACC00803
 ID ACC00803 standard; cDNA; 2463 BP.

XX AC ACC00803;

XX 16-MAY-2003 (first entry)
 DT
 XX
 XX
 XX

Glycine max oil trait related cDNA sequence SEQ ID NO:355.
 DE
 XX
 XX

Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW

CC The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (PP) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, Lip15-like transcription factor activity, calceosin-like activity, ATP citrate lyase activity, SNF1-like activity, and CKC-like transcription factor activity. Also described: (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention

XX Sequence 1824 BP; 573 A; 345 C; 428 G; 478 T; 0 U; 0 Other;

Query Match 20.5%; Score 413.6; DB 7; Length 1824;
Best Local Similarity 79.7%; Pred. No. 1.5e-95;
Matches 488; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 695 CGTTGAAGCTACCGAAGAAACTATTGAGAGTTTGGACAGAGGACGCTCTATATACCG 754
DB 83 CATTGAAACTGCCCCAGAAAGTCCATTGACACTTTGGACAGAGAACTTCTATCTACCG 142
QY 755 CGGTGTTTACAAGGATCGGTGGACAGGAAGATATGAGGCACATTTATGGGATATATAGTTG 814
DB 143 TGGTGTAAACAGGATAGGTGGACGGGAGGTACAGGCTCACCTGTGGGATATATAGTTG 202
QY 815 TAAAGAGAGGCCAAACGCCAAAGAAAGCAAGTTTATTTGGAGTTTATGACAAAGA 874
DB 203 TAGAAGAGAGGACAGACTGCCAAAGAAAGCAAGTTTACTTGGTGGTTATGATAAGA 262
QY 875 AGAAAAAGCAGTACGAGGCTTATGATTTAGCGCACTCAAGTATTTGGGACCACTAC 934
DB 263 AGAAAGGACGCTAGACCTACGATTTGGCAGCACTAAATATCTGGGAAACACCAAC 322
QY 935 TACTAATCTCCCATGAGCGAATATGAAAGAGGTAGAGATGAAAGCAGATGACAAG 994
DB 323 AACAAATTTTCCAATTAGCCACTATGAGAAAGAGTTTGGAAAGATGAAGCACATGACTAG 382
QY 995 GCAGAGATATGTCCTCACTGCCAGAAAGTATGTTCTCTCGTGGTGCATCGAT 1054
DB 383 GCAAGAGTACGTTGCGTCATTGAGAGGAAGAGTAGTGGGTTTCTCGGGTGCATCCAT 442
QY 1055 TTATCGTGGAGTAAACAAGACATCACCAACATGGAAGATGGCAAGCTAGGATAGGAAGAT 1114
DB 443 TTATCGAGGAGTACGAGACACCAACATGGAAGTGGCAAGGATTTGGAAGAT 502
QY 1115 CGCGGTAAACAAGACCTTACTTGGAACTTTGGCACAAGAGAGAGCTCGAGGC 1174
DB 503 TGCTGGCAACAAGGATCTTACTTGGAACTTTAGCACCACCAAGAGAGCGCGAAGC 562
QY 1175 ATACGACATTTGGGCCATCAAAATCAGAGGATTAACCGAGTCACTAACTTGCACATGAA 1234
DB 563 ATATGATGTAGCAGCAATCAAAATCCGAGGACTAAGTGTGTTACAACATTTGACATGAG 522
QY 1235 CAGATCAAGTTAAAGCAATCTTCGAAAGCCCTAGTCTTCTTATTTGGTAGCGCCGCAAA 1294
DB 623 CAGATATGAGCTGAAAGAGCACTATTGAGAGCACCACTTTGCCAATAGTGGTGGTGCAAA 682
QY 1295 ACGTCTCAAGGA 1306
DB 683 GCGTTTGAAGGA 694

RESULT 6
ACCO0801
ID ACC00801 standard; cDNA; 2588 BP.
XX

AC ACC00801;
XX 16-MAY-2003 (first entry)
XX Glycine max oil trait related cDNA sequence SEQ ID NO:351.
XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase; receptor-like protein kinase; mitogen activated protein kinase; oil; Lip15-like transcription factor calceosin; ATP citrate lyase; SNF1; CKC-like transcription factor; antisense inhibition; co-suppression; transgenic plant; gene; ss.
XX Glycine max.
OS WO2003002751-A2.
PN 09-JAN-2003.
PD 27-JUN-2002; 2002WO-US020152.
PF 29-JUN-2001; 2001US-0301913P.
PR (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
PA Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT, Jones TU, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H, Shen B; Tarczynski MC;
PI WPI; 2003-201509/19.
XX P-PSDB; ABR40766.
DR Novel nucleotide fragment encoding polypeptides having receptor-like protein kinase activity, calceosin-like activity, useful for altering oil phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
PS Claim 18; Page 368; 542pp; English.
XX The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (PP) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, Lip15-like transcription factor activity, calceosin-like activity, ATP citrate lyase activity, SNF1-like activity, and CKC-like transcription factor activity. Also described: (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention

SQ Sequence 2588 BP; 863 A; 513 C; 554 G; 658 T; 0 U; 0 Other;

Query Match 20.5%; Score 413.6; DB 7; Length 2588;
Best Local Similarity 79.7%; Pred. No. 1.7e-95;
Matches 488; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 695 CGTTGAAGCTACCGAAGAAACTATTGAGAGTTTGGACAGAGGACGCTATATACCG 754
DB 773 CATTGAACTGCCACCCAGAAAGTCCATTGACATTTTGGACAGAGAACTTCTATCTACCG 832
QY 755 CGGTGTTTACAAGGATCGGTGGACAGGAAGATATGAGGCACATTTATGGGATATAGTTG 814
DB 833 TGGTGTAAACAGGCTAGGTGGACGGGAGGTACGAGGCTCACCTGTGGGATATAGTTG 892
QY 815 TAAAGAGAGGGCCAAACGCCAAAGCAAGCAAGTTTATTTGGAGGTTTATCACAAGA 874
DB 893 TAGAAGAGAGGACAGACTCGCAAGGAAGCAAGTTTACTTGGTGGTTATGATAAGA 952

QY 875 AGAARAGAGCTAGGCTTATGATTTAGCGCAGCTCAAGTATTTGGGGAACACCACTAC 934
DB 953 AGAARAGGAGCTAGAGCTTACGATTTGGCAGCACTAAATATCTGGGGAACACCAAC 1012
QY 935 TACTAATCTCCCATGAGCGAATATCAAAAGAGGTGAGAGATGAACACATGACAAAG 994
DB 1013 AACAAATTTTCCAAATAGCCACTATCAGAAAGAGTTGGAAGAAATGAAGCACATGACTAG 1072
QY 995 GCAAGAGTATGTCCTCTAGCGGAGGAAAGAGTGTCTCTCGGTGTCATCGAT 1054
DB 1073 GCAAGAGTATGTCCTCTAGCGGAGGAAAGAGTGTCTCTCGGTGTCATCGAT 1132
QY 1055 TTATCGTGGAGTAACAGACATCACCAACATGGAAGATGCGAAGCTAGATAGGAGAGT 1114
DB 1133 TTATCGAGAGTACGAGACACCAACCAACATGGAAGTGGCAAGGAGATTTGGAAGT 1192
QY 1115 CGCGGTAAACAAGACTCTACTTGGGAATTTTGGCACAACAGAGAGCTGCGAGGC 1174
DB 1193 TGTGCGCAACAGGATCTTACTTGGGAATTTTAGCACCACCAAGAGAGCGAGCGAAGC 1252
QY 1175 ATACGACATTCGCGGCATCAAAATTCAGAGGATTAACCGGAGTGACTTAACTTCGACATGAA 1234
DB 1253 ATATGATGAGCAGCAATCAAAATTCGAGGACTAAGTGTCTTTACAAACTTTGACATGAG 1312
QY 1235 CAGATACAGGTTAAAGCAATCTCGAAAGCCCTAGTCTTCTCTATTTGGTAGCGCCGCAAA 1294
DB 1313 CAGATATGAGTGAAGAGCATACTTGAGAGCACCACTTTGCCAATAGGTGGTGTGCAAA 1372
QY 1295 ACCTCTCAAGGA 1306
DB 1373 GCGTTTGAAGGA 1384

RESULT 7

AAC83248
ID AAC83248 standard; DNA; 1749 BP.
XX AAC83248;
AC AAC83248;
DT 14-MAR-2001 (first entry)
XX Cellulose synthase promoter sequence SEQ ID 3.
DE Cellulose synthase gene expressed during deposition of secondary cell
KW Cellulose synthase; lignin; secondary cell wall construction; wood pulp;
KW transgenic plant; paper manufacture; promoter; ds.
XX Unidentified.
OS WO200070058-A2.
XX
XX
XX 23-NOV-2000.
XX
XX 17-MAY-2000; 2000WO-GB001890.
XX
XX 18-MAY-1999; 99GB-00011379.
XX (UYMA-) UNIV VICTORIA MANCHESTER.
XX
XX Turner S, Taylor N;
XX WPI; 2001-041015/05.
XX
XX Cellulose synthase gene expressed during deposition of secondary cell
PT walls in lignin-containing cells; useful for modulating expression of
PT enzymes involved in synthesis of plant cell walls and to produce
PT transgenic plants.
XX
XX Claim 10; Page 30-31; 49pp; English.
XX
XX This invention relates to a cellulose synthase gene expressed during the
CC deposition of secondary cell walls in cells containing lignin. The
CC cellulose synthase gene is useful for regulating the expression of genes

on cellulose

CC specifically during secondary cell deposition in lignin containing cells.
CC it can be used to modify the structure and cellulose content of plant
CC secondary cell walls and to produce altered plant phenotypes specific to
CC the needs of a particular industry such as in reducing the lignin of wood
CC pulp for paper manufacturing. A construct containing a cellulose synthase
CC promoter sequence and a gene of interest may be used in a method for the
CC production of the product of the gene of interest in a host cell that
CC produces lignin, where the product is produced only during secondary cell
CC wall synthesis. The present sequence represents a cellulose synthase
CC promoter which can be used in the invention for the production of
CC transgenic plants expressing an exogenous gene during secondary cell wall
CC deposition in cells containing lignin

XX SQ Sequence 1749 BP; 551 A; 290 C; 354 G; 554 T; 0 U; 0 Other;

Query Match 19.8%; Score 399.2; DB 4; Length 1749;
Best Local Similarity 77.8%; Pred. No. 7.1e-92;

Matches 502; Conservative 0; Mismatches 123; Indels 49; Gaps 8;

QY 1257 CTCGAAGCCCTAGTCTTCTTATTTGGTAGCGCCCAAAAGCTCTCAAGGAGCTAACCGT 1316
DB 1 CTCGAGAGCCCGAGTCTACCTATTGGTAGTTCTCGAAACGCTCTCAAGGAGCTTAAACAAT 60
QY 1317 CCGGTTCCAAAGTATGATGATCAGTAATAACGTTTCAGAGAGTGAGAAATAGTCTAGC 1376
DB 61 CCGGTTCCAGCT---ATGATGATTAATAACGTTTCAGAGAGTGCAATTAATGTTAGC 117
QY 1377 GGTGGCAAAAACGCTGCGGTTTCAGATCATCAGGAGTAGATTTGAGCTTATTCACCAAA 1436
DB 118 GGTGGCAAAAACACTGCGTTTCAG---CATCAGGAAATGGATTTGAGCTTATTCAGCAA 174
QY 1437 CATCAAGAGAGGTACAAATGTTTATTAACAATGAGGAAACTTGTCTCGGAGAGTCT 1496
DB 175 CAGCAGGAGAGGTAGCTGG---TTATTAACAATGAGGAAACTTGTCTCGGAGAGTACT 231
QY 1497 AGGCTTGTCTTCAAAACAGAGAGTATCAACACCATTTCTTGAGCAACACAGAGCTC 1556
DB 232 AGGCTTGTCTTCAAAACAGAGAGGAGCAACAAACACTTCTTGAGAAACTCGCGAGTCA 291
QY 1557 ATGACTAATATCGATCATCAAGTCTGTTTGGAGTATCGGTACTGTTTGTGGAAT 1616
DB 292 ATGACTAATATGATCATCATAGCTCTACCTCTGATGATTTCTGTTACCGTTTGTGGAAT 351
QY 1617 GTTGTGCTTATGTTGTTTATCAAGGATTTGAGCC---CGGTTAACTGC 1664
DB 352 GTTGTGCTTATGTTGTTTATCAAGGATTTGCAATCCCTGTTGGAACATCGTTAATTAC 411
QY 1665 GATGCTAGCTGCTAGTGTGATTAACCGCAAGAAACCATTTACTTTGCTCAG 1724
DB 412 GATGCTTACTGCTGCTGAGATTTGCTTCAACGCAAGAAATCATTTACTATGCTCAG 471
QY 1725 CAG-----CAGCAGACCCAGCTGCGAGTGAGATTTCCCGCGCAATGACG 1775
DB 472 CATCAGCAACACAGCAGATTCAGAGCTCGCGGAGAGATTTCCGCTGCAATTTGCG 531
QY 1776 AATAAGTTGGCTTAATATATATTAATCAATCGGAGAGCTGTTGAGAGATTTGCTCCAA 1835
DB 532 AATAACCATAGCTTAACATGATTTTCAAGGAGAGCTGTTGAGAGAGGCGCTCCAAAG 591
QY 1836 TTTACAGTTTGAACGACAATTAAGAAATAGTTTAAAGATCTTTAGTTATATGCTGTTGT 1895
DB 592 TTTTTCAGTTTGAACGACACTTAGAAAAATAGTAAAGATCTTT-----TAGT 640
QY 1896 TGTGCTGCTGGTGAACAGTGTGATCTTTGATATGTTTCTTTCTCTCTCTCTCTCTCT 1955
DB 641 TGTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 693
QY 1956 CTGTTGTTAATTTCTTAAGACATTTT---TTAGTTTCCATGTTGATTAATTTT 2008
DB 694 TTGGGTAATTTCTTATAAATTTTTCATAGTTTTCGATTTATTTGGATAAATTT 747

RESULT 8

ACC00806
 ID ACC00806 standard; cDNA; 1990 BP.
 XX
 AC ACC00806;
 DT
 DT 16-MAY-2003 (first entry)
 DE Glycine max oil trait related cDNA sequence SEQ ID NO:361.
 XX
 XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ARP citrate lyase; SNF1;
 KW CK-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX
 OS Glycine max.
 XX
 XX WO2003002751-A2.
 PN
 XX
 XX 09-JAN-2003.
 PD
 XX
 XX 27-JUN-2002; 2002WO-US020152.
 PF
 XX
 XX 29-JUN-2001; 2001US-0301913P.
 PR
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 PA
 XX
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczyński MC;
 XX
 XX WPI; 2003-201509/19.
 DR
 DR P-PSDB; ABR40771.
 XX
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 PS Claim 18; Page 381-382; 542pp; English.
 XX
 CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ARP citrate lyase activity, SNF1-like
 CC activity and CK-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 XX Sequence 1990 BP; 628 A; 391 C; 459 G; 512 T; 0 U; 0 Other;
 SQ
 Query Match 19.2%; Score 385.8; DB 7; Length 1990;
 Best Local Similarity 76.9%; Pred. No. 2.1e-88;
 Matches 484; Conservative 0; Mismatches 142; Indels 3; Gaps 1;
 XX
 QY 692 TAGCGTTGACCTACCGAGAAACTATTGAGAGTTTGGACAGAGGACGCTATATA 751
 |||
 DB 560 TACTGTTGAGTTGACCTACCGAGAACTTTGGATACATTGGGACAGAGAACTCCATATA 619
 |||
 QY 752 CCGCGGTTTACAGGATCGGTGACAGGAGATATGAGGCACATTATTGGGATAATAG 811
 |||
 DB 620 TCGTGGAGTAACCTGACATAGATGGAAGTATGAAGTTCATCTTGGGATAATAG 679
 |||

812 TTGTAAGAGAGAGCGCCAAACCGCAAGAGAGAGTATTTATTTGGGAGGTTATGACAA 871
 |||
 DB 680 CTGTAGAGGGAAGCCCAATCAAGAAAGGAGCCCAAGTTTATTTGGGTGATATGATAA 739
 |||
 QY 872 AGAAGAAAAAGCAGCTTAGGGCTTATGATTTAGCCGCACTCAAGTATTCGGGAACACCAC 931
 |||
 DB 740 AGAAGAGAAAGCAGCTTAGAGCTTATGATTTAGCTCAGTGAAGTACTGGGGACATCCAC 799
 |||
 QY 932 TACTTACTTCCCATGAGCGAATATGAAAAGAGGTAGAGATGAAGCAATGAC 991
 |||
 DB 800 CACTACCACTTTCCAAATAGCACTATGAGAAAGGAATTTGATGAATGAACACATGAC 859
 |||
 QY 992 AAGGCAAGAGTAGTTGTCCTCACTGCGCAGGAAAAAGTAGTGTCTCTCTCGTGTGATC 1051
 |||
 DB 860 GAGCAAGAATTTGTTGCGCCATTAGAAGGAAGAGCAGTGTCTTCCAGGGGTGATC 919
 |||
 QY 1052 GATTTATCTGAGTAACAAGACATCAACAATGGAAGATGGCAAGCTAGATAGGAAG 1111
 |||
 DB 920 AATGTATCTGTGAGTTTACAAGGCATCCCAACACGGAAGATGGCAAGCAAGATTGGCAG 979
 |||
 QY 1112 AGTCGCGGTAAACAAAGACCTCTACTTGGGAACCTTTTGGCACACAAAGAAAGCTGCAGA 1171
 |||
 DB 980 AGTTGCAGAAACAAAGATCTTTACTTGGGAACCTTTCACTAGTGAAGAGGCTGCAGA 1039
 |||
 QY 1172 GGCATACGACATTTGGGCGCATCAAAATTCAGAGGATTAACCGAGTGAATCTTGCACAT 1231
 |||
 DB 1040 AGCATACGACATAGCAGCGATAAAGTTTCAGAGGTCTCAACGCTGTCACAAACTTTGACAT 1099
 |||
 QY 1232 GAACAGATACAACGTTAAAGCAATCTCGAAAGCCCTAGTCTTCTCTAT---TGGTAGGCG 1288
 |||
 DB 1100 GAGCGCTACGACGTGAAGGCCATTTCTTGAAGGCAACACTCTCCCAATAGGAGGAGGCG 1159
 |||
 QY 1289 CGCAAAACGTTCAAGGAGGCTTAACCGTC 1317
 |||
 DB 1160 TCGAAAGGCTGTGAAGAAAGCTCAAGCTC 1188
 |||

RESULT 9
 ADD30878
 ID ADD30878 standard; cDNA; 1818 BP.
 XX
 AC ADD30878;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Plant yield-related polynucleotide clone G1793.
 XX
 KW ds; transcription factor; transgenic plant; growth rate; senescence;
 KW seed germination rate; plant vigor; seedling vigor.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2003013227-A2.
 XX
 PD 20-FEB-2003.
 PF
 XX 09-AUG-2002; 2002WO-US025805.
 XX
 XX 09-AUG-2001; 2001US-0310847P.
 PR 19-NOV-2001; 2001US-0336049P.
 PR 11-DEC-2001; 2001US-0338692P.
 PR 14-JUN-2002; 2002US-00171468.
 XX
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX
 XX Ratcliffe O, Rischmann JL, Adam LJ, Dubell AT, Heard JE;
 PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
 PI Brown PE;
 XX
 XX WPI; 2003-248221/24.
 DR P-PSDB; ADD30879.
 XX
 PT New plant transcription factor polynucleotides and polypeptides, useful

PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 907; 454pp; English.
XX
CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcription
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the cDNAs of the invention.
XX
SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;
Query Match 18.7%; Score 377.4; DB 9; Length 1818;
Best Local Similarity 76.7%; Pred. No. 2.8e-86;
Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 702 GCTACACCGAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
DB 554 GCCACGCCAAGACGTGCATTGGACACTTCGGACAAAGCAACCTCGATCTATCGTGTGTC 613
QY 762 ACAAGGCATCGGTGGACAGGAAGATATGAGCACTTATGGGATATATAGTTGTAAGA 821
DB 614 ACAAGACATCGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 673
QY 822 GAAGGCCAAACCGCAAGGAAGACAAGTTTATTTGGAGGTTTATGACAAAGAAAGAAA 881
DB 674 GAAGGCCACTCTAGGAAGGAAGACAAGTTTACTTGGTGGATATGACAAAGAAAGTAA 733
QY 882 GCAGCTAGGCTTATGATTATGGCGACTCAAGTATTGGGGAACCACTACTACTAAC 941
DB 734 GCAGCAAGATCATATGATATGATGCTGCTTATGAGTCTTCACTACTACTAT 793
QY 942 TTCCCATGACGAATATGAAAGAGGTAGAGAGATGAAGACATGACAAAGCAAGAG 1001
DB 794 TTCCCATTACAACTACGAGAAAGATGAGGAAATGAAGACATGACGACAAAGAG 853
QY 1002 TATGTTGCTTCACTGCGCAGGAAGTATGTTTCTCTGCTGTCATCGATTATCGT 1061
DB 854 TCGTGGCTGCTTATGAGGAAAGATGAGTGGATTTTCGAGAGCGCTTCGATATCGA 913
QY 1062 GGAGTAAACAGACATCAACATGGAAGATGGCAAGTGGCAAGCAAGGATCGGCCGCGGA 973
DB 914 GGAGTTACAGGCATCAACATGGAAGATGGCAAGCAAGGATCGGCCGCGCGGA 973
QY 1122 AACAAAGACCTTACTTGGGAACTTTGGACACAAAGAAAGCTCGAGAGCATAGC 1181
DB 974 AACAAAGACCTTACTTGGGAACTTTAGCACTGAGGAAAGCAAGCAAGGTTAGCAT 1033
QY 1182 ATTCCGGCCATCAATTCAGAGGATTAACCGCAGTACTAATTCGACATCAACAGATAC 1241
DB 1034 ATAGCTGCAATAAGTTTAGAGGCTTAATGCACTGACCACTTCGAGATCAACCGGTAC 1093
QY 1242 AACGTTAAAGCAATCTCGAAAGCCCTAGTCTTCCCTAT---TGTAGCGCGCAAAAGCT 1298
DB 1094 GACGTGAAGGCAATTCATAGAGATAGCACTCTTCCCATCGAGGAGCGCAGCTAAACGG 1153
QY 1299 CTCAGGAGGCTAACCGTCT 1317
DB 1154 CTCAGGAGGCTCAAGCTC 1172
RESULT 10
ADD30822
ID ADD30822 standard; cDNA, 1818 BP.

XX
AC ADD30822;
XX
DT 15-JAN-2004 (first entry)
XX
DE Plant yield-related polynucleotide clone G1793.
XX
KW ds; transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.
XX
OS Arabidopsis thaliana.
XX
FN WO2003013227-A2.
XX
PD 20-FEB-2003.
XX
PF 09-AUG-2002; 2002WO-US025805.
XX
PR 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P.
PR 14-JUN-2002; 2002US-00171468.
XX
(MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PA Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI Brown FE;
XX
DR WPI; 2003-248221/24.
DR P-PSDB; ADD30823.
XX
PT New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 851; 454pp; English.
XX
CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcription
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the cDNAs of the invention.
XX
SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;
Query Match 18.7%; Score 377.4; DB 9; Length 1818;
Best Local Similarity 76.7%; Pred. No. 2.8e-86;
Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 702 GCTACACCGAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
DB 554 GCCACGCCAAGACGTGCATTGGACACTTCGGACAAAGCAACCTCGATCTATCGTGTGTC 613
QY 762 ACAAGGCATCGGTGGACAGGAAGATATGAGCACTTATGGGATATATAGTTGTAAGA 821
DB 614 ACAAGACATCGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 673
QY 822 GAAGGCCAAACCGCAAGGAAGACAAGTTTATTTGGAGGTTTATGACAAAGAAAGAAA 881
DB 674 GAAGGCCACTCTAGGAAGGAAGACAAGTTTACTTGGTGGATATGACAAAGAAAGTAA 733
QY 882 GCAGCTAGGCTTATGATTATGGCGACTCAAGTATTGGGGAACCACTACTACTAAC 941
DB 734 GCAGCAAGATCATATGATATGATGCTGCTTATGAGTCTTCACTACTACTAT 793

KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX
 OS Cossypium hirsutum.
 XX
 XX WO2003002751-A2.
 XX
 PD 09-JAN-2003.
 XX
 XX 27-JUN-2002; 2002WO-US020152.
 PF
 XX 29-JUN-2001; 2001US-0301913P.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TU, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX
 DR WPI; 2003-201509/19.
 DR F-PSDB; ABR40854.
 XX
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 PS Claim 18; Page 503; 542pp; English.
 XX
 CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity, and CKC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 1935 BP; 576 A; 436 C; 416 G; 507 T; 0 U; 0 Other;

Query Match 17.7%; Score 357.4; DB 7; Length 1935;
 Best Local Similarity 73.8%; Pred. No. 48-81;
 Matches 454; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 658 ACAACGTTGTTGCCAAGGGAAGACTATTGATGATAGCGTTGAAGCTACACCGAAGAAA 717
 DB 442 AATCGATATTCAACAAACACGCAACAGCACCTATAGCTGAGCTGCACCCAAAAGAT 501
 QY 718 CTATTGAGATTGTCACAGAGAGCTCTATATACCGCGTGTACAGGCATCGGTGA 777
 DB 502 CTGTTGATCTTTTGGCCAACTACCTCGATCTACCGGGGTGTCCACCGCATAGATGA 561
 QY 778 CAGGAATATGCGCACATTTATGGGATATAGTTCTTAAAGAGAGGCGCAACCGCA 837
 DB 562 CTGAAGATATGAAGCTCACTTGTGGGACATAGTTGTCAGAGAGAGGCGCAAGTAGGA 621
 QY 838 AAGGAAGACAGTTTATTGGGAGGTTATGCAAGAGAAAAGAGAGCTAGGCTTATG 897
 DB 622 AAGGAAGCAAGTTTATTGGGCTGCTATGACAGGAGATATAAGCTGCAAGAGCTTATG 681
 QY 898 ATTTAGCGGCACTCAAGTATTGGGGAACCCACTACTACTACTTCCCATGAGCGAAT 957

DB 682 ATCTTGGGCTCTCAAGTACTGGGGTCCGACCACTACTACCAACTTTCGATTTCCAAT 741
 QY 958 ATGAAAAGAGGTAGAAAGATGAAGCATGACAAAGGCAAGAGATATGTTCCCTCACTGC 1017
 DB 742 ATGAGAAGAGCTCGAAGAGATGAGAACATGACTAGGCAAGAGTTGTTGCTCTCTCC 801
 QY 1018 GCAGGAAAAGTAGTGGTTCTCTCGTGGTCATCGATTTATCGTGGAGTAACAGACATC 1077
 DB 802 GGAGGAAAAGTAGTGGATTTTCTAGGGGTGCTTCAATTTACAGAGGAGTGACAAAGCATC 861
 QY 1078 ACCAATCATGGAAGTGGCAAGCTAGGATAGGAAGAGTGCGCCGTAAACAAAGAGCTTACT 1137
 DB 862 ATCAATCATGTTGGTGGCAAGCAAGATTGGAGAGATTGCGAGCAACAAGATCTCTATC 921
 QY 1138 TGGGAACCTTTTGGCACACAAAGAAAGACTCGCAGGCGATACGACATTCGGCCATCAAT 1197
 DB 922 TTGCACATTTAGCACCCAAAGAAAGAGCTGAAGCCTATGATATTGCGACCAATCAAGT 981
 QY 1198 TCAGAGATTAAACCGGAGTGAATTCGACATGAACAGATACCAAGTTTAAAGCAATCC 1257
 DB 982 TTAGAGGCTTTAAACCGGTTTACAAATTTTGATATGAGCCCTAGCTAGTAAACAGCAT 1041
 QY 1258 TCGAAAGCCCTAGTC 1272
 DB 1042 CAACAGCAATCTTC 1056
 XX
 AC ACC00804;
 XX
 DT 16-MAY-2003 (first entry)
 XX
 DE Glycine max oil trait related cDNA sequence SEQ ID NO:357.
 XX
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX
 OS Glycine max.
 XX
 XX WO2003002751-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 27-JUN-2002; 2002WO-US020152.
 XX
 XX 29-JUN-2001; 2001US-0301913P.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TU, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX
 XX WPI; 2003-201509/19.
 DR F-PSDB; ABR40769.
 XX
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 PS Claim 18; Page 377; 542pp; English.
 XX
 CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity, and CKC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 1935 BP; 576 A; 436 C; 416 G; 507 T; 0 U; 0 Other;

CC activated protein (MAP)-kinase activity, LIP15-like transcription factor
 CC activity, caloesin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CKC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (III), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC inhibition or co-suppression in a plant such as corn, soybean, wheat, rice,
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 2168 BP; 622 A; 570 C; 432 G; 544 T; 0 U; 0 Other;

Query Match 17.5%; Score 352; DB 7; Length 2168;
 Best Local Similarity 75.2%; Pred. No. 1e-79;
 Matches 439; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
 QY 700 AAGCTACACCGAAGAAACTATTGAGAGTTTGGACAGAGACGCTCTATATACCGCGGTG 759
 Db 517 AACCATCCCCCAGAAACCTCGACACCTTCGGGCAACGACCTCATCTACCGCGGCG 576
 QY 760 TTACAGCGCTCGGTGGACAGAAATATAGGACACATTTATGGATAATAGTTGTAAAA 819
 Db 577 TGACCCGACATAGATGGACTGGGAGATACGAAGCTCATCTATGGCAATAGTTGCAGAA 636
 QY 820 GAGAGGCGCAACCGCGCAAGAGAGACAGTTTATTTGGGAGGTTATGACAAAGAGAAA 879
 Db 637 GGGAGGTCAAGCAGGAGAGAGGAGCAAGTTACTTGGTGGTTATGACAGAGGATA 696
 QY 880 AAGCAGTAGGCTTATGATTTATCCGCACTCAAGTATGGGAAACACCACTACTACTA 939
 Db 697 AGGCAGCCAGAGCTTATGATCTCGAGCTCTCAAGTACTGGGTCCAACTACCACTA 756
 QY 940 ACTTCCCATGAGCGAATATGAAAGAGGTAGAGAGATCAACACATGACAGGCAAG 999
 Db 757 ACTTCTCTATTTCACTATGAGAGGAATGAGGAGATGAAGACATGACTAGGCAAG 816
 QY 1000 AGTATGTTGCTCTACTCGCAGGAAAAGTAGTGTGTTTCTCTCGTGTGTCATCGATTATC 1059
 Db 817 AGTTGTTGTTCTCTCTCGTAGGAAGAGCAGTGTGTTCTCTAGAGGGCTCTATATACA 876
 QY 1060 GTGGAGTAACAAGACATCCACATCGAAGATGGCAAGTAGTAGTAGGAGTGGCGG 1119
 Db 877 GAGGAGTAACGAGACACCAACGATGGCGGATGGCAGGCGAGAAATAGGCAAGTTGCCG 936
 QY 1120 GTAAACAAGACCTTACTTGGGAACCTTTGGGCACACAAGAAAGAGCTGACAGGGCATACG 1179
 Db 937 GAAACAAGACCTTACTTGGGACCTTTCAGACCCCAAGAGAGCTGCTGAGGCTATG 996
 QY 1180 ACATTGGGCAATCAATTCAGAGGATTAACCGAGTACTACTTCCGACATGAACAGAT 1239
 Db 997 ACATTGCTGTATCAATTCAGGGGATTAATCGATCAACAACTTTGACATGAGTCGCT 1056
 QY 1240 ACAACGTTAAAGCAATCCTCGAAAGCCCTAGTCTTCTCTATTGTT 1283
 Db 1057 ACGAGGTGAGAGCATTTGCAATAGTACTTCTCTATTGTTGTT 1100

RESULT 14

ACC00805
 ID ACC00805 standard; cDNA; 1290 BP.

XX
 AC ACC00805;

XX
 DT 16-MAY-2003 (first entry)

DE Glycine max oil trait related cDNA sequence SEQ ID NO:359.

XX
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;

KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP15-like transcription factor caloesin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX Glycine max.
 XX WO2003002751-A2.
 XX
 XX 09-JAN-2003.
 PD
 XX
 XX 27-JUN-2002; 2002WO-US020152.
 XX
 XX 29-JUN-2001; 2001US-0301913P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX
 XX WPI; 2003-201509/19.
 XX P-PSDB; ABR40770.
 XX
 PT Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caloesin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 PS Claim 18; Page 379-380; 542pp; English.
 XX
 CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP15-like transcription factor
 CC activity, caloesin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CKC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (III), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 1290 BP; 414 A; 231 C; 321 G; 324 T; 0 U; 0 Other;

Query Match 17.3%; Score 347.8; DB 7; Length 1290;
 Best Local Similarity 73.8%; Pred. No. 9.7e-75;
 Matches 442; Conservative 0; Mismatches 157; Indels 0; Gaps 0;
 QY 699 GAAGTACACCGAAGAAACTATTGAGAGTTTGGACAGAGGACGCTCTATATACCGCGGT 758
 Db 117 GAACAGCTTCTCTCGCAATCTATTGACACTTTGGGCAAGAGACATCTCATATCGTGGT 176
 QY 759 GTTCAAGGCGATCGGTGGACAGGAAGATATGAGCACAATTATGGGATAATAGTTGAAA 818
 Db 177 GTTCAAGGCGATAGATGAGTGAAGATATGAGCCCACTTTGGGATAATAGTTGCGA 236
 QY 819 AGAGAAGGCCAAACGCGCAAGAGAGCAAGTTTATTTGGGAGGTTATGACAGAGAGAA 878
 Db 237 AAGGAGGGGCAACCAAGGAGGAGGCAAGTTTACCTTGTGTTAATAGGAGAGAA 296
 QY 879 AAAGCAGCTAGGGCTTATGATTATAGCCGCACTCAAGATTATTTGGGAACCACTACTACT 938
 Db 297 AAAGCAGCAAAAGGCTTACGACTTAGTGTGCACTCAAGTATTGGGTCCAACTACATA 356
 QY 939 AACTTCCCATGAGCGAATATGAAAGAGGTAGAGAGATGAAGCAGATGACACATGACAGGCA 998

357 AATTTCCCTTTAAGCACCTTATGAGAAGGAACCTTGAAGAGATGAAGACATGACGAGCAA 416
 999 GAGTATGTTGCTCACTCGGAGGAAGTATGTTCTCTCTGTTGTCATCGATTAT 1058
 417 GAATTTGTCGCAATTAAGGAAGGAAGAGTGAATTTTCAAGAGGGGCACTGTGTAT 476
 1059 CTTGAGTAACAAGACATCAACCAATGGAAGATGGCAAGCTAGGATAGGAAGTGCCT 1118
 477 AGAGTGTGACAGACACCATCAACATGGAAGTGGCAGGCCCCGAATAGGAAGGTTGCA 536
 1119 GGTACAAAGACCTTACTTGGGAACCTTTGGCACACAGGAAGCTGAGGACATAC 1178
 537 GGAACAAGAGACTTGTATCTTGTATCTTGTATCTTGTATCTTGTATCTTGTATCT 596
 1179 GACATTGGGGCCATCAATTTAGAGGATTAACCGCAGTGAATTAATTCGACATGAACAGA 1238
 597 GATATAGCAGCANTCAAGTTAGAGGAACAGTGTGTGACCAACTTTGACATAAGTAGG 656
 1239 TACACAGTTAAAGCAATCTCGAAGAGCCCTAGTCTTCTTATGTTAGGCGCGCAAAAG 1297
 657 TATGATGTGAAGAGATCTGCTCAAGCTCCACTCTTATTGCGAGGAGATCTTGCAGAAAG 715

RESULT 15

ACC00850
 ID ACC00850 standard; cDNA; 1954 BP.

XX
 AC ACC00850;

XX
 16-MAY-2003 (first entry)

DE Glycine max oil trait related cDNA sequence SEQ ID NO:490.

XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.

XX Glycine max.

XX WO2003002751-A2.

XX 09-JAN-2003.

XX 27-JUN-2002; 2002WO-US020152.

XX 29-JUN-2001; 2001US-0301913P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PION-) PIONEER HI-BRED INT INC.

XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;

XX WPI; 2003-201509/19.
 XX P-PSDB; ABR40856.

XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 XX Claim 18; Page 508-509; 542pp; English.

XX The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein kinase (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CKC-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)

CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00850 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention

XX
 SQ Sequence 1954 BP; 582 A; 379 C; 468 G; 521 T; 0 U; 4 Other;

Query Match 17.3%; Score 347.8; DB 7; Length 1954;
 Best Local Similarity 73.8%; Pred. No. 1.2e-78;
 Matches 442; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 699 GAAGCTACACCGAAGAAACTATTGAGAGTTTGGACAGAGGACGTTTATACCGCGGT 758
 DB 781 GAACGAGTCTCTCGCAATCTATTGACATTTTGGGCAAGAACATCTCAATATCGTGT 840
 QY 759 GTTACAGGCGATCGTGGACAGGAAGATATGAGGCACATTTATGGGATATAGTTGAAA 818
 DB 841 GTTACAGGCGATAGTGGACTGGAAGATATGAGGCCCATTTTGGGATATATAGTTGAGA 900
 QY 819 AGAGAGGCCAAACCGCGCAAGAAAGAACAAAGTTTATTTGGAGGTTATGCAAGAGAA 878
 DB 901 AAGGAGCGGCAACAGGAAGGAAGCAAGTTTACCTGCTGTTATGATAGGAAGAA 960
 QY 879 AAGCAGCTAGGCTTATGATTTAGCCGCACTCAAGTATTTGGGACACCACCTACTACT 938
 DB 961 AAGCAGCAAGGCTTACGACTTAGCTGCACTCAAGTATTTGGGGTCCCAACACTCACATA 1020
 QY 939 AACTTCCCATGAGCGAATATGAAAAGAGTGAAGAGATGAAGACATGACAAAGGCAA 998
 DB 1021 AATTTCCCTTTAGCACTTATGAGAGAACTTGAAGAGATGAAGACATGACCAAGCAA 1080
 QY 999 GAGTATGTTGCTTCACTCGCGCAGGAAAGTAGTGGTTTCTCTGTTGTCATCGATTAT 1058
 DB 1081 GAATTTGTCGCAATTTAAGAAGGAGAGCAGTGGATTTTCAAGAGGGGCACTCTGTAT 1140
 QY 1059 CGTGGAGTACAGACATCCACATGGAAGATGCGAAGCTAGATAGGAAGAGTGCCT 1118
 DB 1141 AGAGGTGACAGACACCATCAACATGGAAGGTTGGCAGGCCCGCAATAGGAAGGTTGCA 1200
 QY 1119 GGTAAACAAAGACCTTCTACTTTGGGAACTTTTGGCACACAAGAAAGCTGCGAGGCTATC 1178
 DB 1201 GGAAACAAGGACTTGTATCTTGTATCTTTAGCACACAGAAAGAGCTGCTGAAGCATAT 1260
 QY 1179 GACATTTGGGGCCATCAATTTGAGAGGATTAACCGCAGTGTACTTCTGACATGACAGA 1238
 DB 1261 GATATAGCAGCAATCAAGTTTCAGAGGAACAAAGTCTGTGACCAACTTTGACATAGTAGG 1320
 QY 1239 TACACGTTAAAGCAATCTCTGAAAGGCCCTAGTCTTCTTATTTGGTAGCGCGCAAAAG 1297
 DB 1321 TATGATGTGAAGAGATCTGCTCAAGCTCCACTCTTATTGCGAGGAGATCTTGCAGAAAG 1379

Search completed: July 19, 2004, 06:39:44
 Job time : 828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 00:07:47 ; Search time 7901 Seconds
(without alignments)
11048.332 Million cell updates/sec

Title: US-09-980-364-1
Perfect score: 2014
Sequence: 1 gtcacatctctctcttaag.....gttgataaatttcagact 2014

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vt.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_in.*
- 18: em_mu.*
- 19: em_mu.*
- 20: em_mu.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vt.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_ror.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	100.0	2014	2014	6	AX058687 Sequence
2	100.0	2014	2014	8	AF317904 Brassica
3	92.8	1888.6	1992	8	AF317905 Brassica
4	92.8	1888.6	2011	6	AX058689 Sequence
5	726	726	4873	6	AX058681 Sequence
6	36.0	36.0	4873	8	AF317906 Brassica
7	504.8	504.8	33563	8	ATT10B6
8	502.6	502.6	25.0	6	AX058692 Sequence
9	502.6	502.6	25.0	8	AF317907 Arabidops
10	399.2	399.2	19.8	6	AX048247 Sequence
11	339.8	339.8	16.9	8	AY461432 Nicotiana
12	330.6	330.6	16.4	8	AY117207 Arabidops
13	330.6	330.6	16.4	8	ATU41339 Arabidops
14	330.6	330.6	16.4	8	ATU44028 Arabidops
15	330.6	330.6	16.4	8	AY080706 Arabidops
16	330.6	330.6	16.4	6	BD274516 Methods f
17	330.6	330.6	16.4	6	AR316367 Sequence
18	330.6	330.6	16.4	6	AR427901 Sequence
19	330.6	330.6	16.4	8	U40256 Arabidops
20	327.2	327.2	16.2	6	AX555218 Sequence
21	324.8	324.8	16.1	6	BD274518 Methods f
22	324.8	324.8	16.1	6	AR316369 Sequence
23	324.8	324.8	16.1	6	AR427903 Sequence
24	323.8	323.8	16.1	8	AK111891 Oryza sat
25	317.8	317.8	15.8	8	AK101959 Oryza sat
26	315	315	15.6	6	AX555220 Sequence
27	315	315	2510	8	AK106306 Oryza sat
28	308.2	308.2	15.3	6	AX555216 Sequence
29	299.8	299.8	14.9	8	ZMWHCF1
30	299	299	14.8	8	AY062179 Oryza sat
31	269.2	269.2	13.4	8	AK109839 Oryza sat
32	255.6	255.6	12.7	6	AX555223 Sequence
33	250.2	250.2	12.4	8	AY133880 Arabidops
34	250.2	250.2	12.4	8	AY045915 Arabidops
35	246.2	246.2	12.2	8	AK109848 Arabidops
36	228.4	228.4	11.3	8	AK111617 Oryza sat
37	214	214	10.6	8	AK106769 Oryza sat
38	205.2	205.2	10.2	8	AK112088 Oryza sat
39	192.8	192.8	9.6	8	ATH580379 Arabidops
40	192.4	192.4	9.6	8	AK117300 Arabidops
41	190.6	190.6	9.5	8	AK112070 Oryza sat
42	161.2	161.2	8.0	8	BT008926 Triticum
43	152.2	152.2	7.6	8	AY088118 Arabidops
44	145	145	7.2	8	AF325506 Pisum sat
45	142.8	142.8	7.1	6	AX505843 Sequence

ALIGNMENTS

RESULT 1
AX058687 LOCUS AX058687 2014 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 1 from Patent WO0075330.
ACCESSION AX058687
VERSION AX058687.1 GI:12311028
KEYWORDS
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1
AUTHORS Boutilier K., Ouellet, T., Custers, J., Hattori, J., Miki, B. and van
lookeren Campagne, M.

intron 1846..2298
intron 2720..2952
intron 3036..3160
intron 3170..3214
intron 3404..3553
intron 3628..3797
intron 3849..3961
intron 4039..4148
misc_feature 4856..4858
/note="stop codon"
ORIGIN
Query Match 36.0%; Score 726; DB 6; Length 4873;
Best Local Similarity 100.0%; Pred. No. 1.3e-162;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1149 GGCACACAGAAGAGCTGAGAGGATACGACATTCGGCCCATCAAAATTCAGAGGATTA 1208
Db 4148 GGCACACAGAAGAGCTGAGAGGATACGACATTCGGCCCATCAAAATTCAGAGGATTA 4207
QY 1209 ACCGAGTGAATTAATTCAGATGAAACAGATCAACGTTAAAGCAATCTCGAAGCCCT 1268
Db 4208 ACCGAGTGAATTAATTCAGATGAAACAGATCAACGTTAAAGCAATCTCGAAGCCCT 4267
QY 1269 AGTCTTCCTATTGTTAGCGCGCAGAAACGTCCTCAAGAGCTCAACGTCGCGTTCACAGT 1328
Db 4268 AGTCTTCCTATTGTTAGCGCGCAGAAACGTCCTCAAGAGCTCAACGTCGCGTTCACAGT 4327
QY 1329 ATGATGATGATCATGATTAATTAACGTTTCAGAGAGTGAGATAGTCTAGCGGTTGGCAAAAC 1388
Db 4328 ATGATGATGATCATGATTAATTAACGTTTCAGAGAGTGAGATAGTCTAGCGGTTGGCAAAAC 4387
QY 1389 GCTGCGGTTGAGTATCATCAGGAGTAGATTTGAGTCTTATTCGACCAACATCAAGAGAGG 1448
Db 4388 GCTGCGGTTGAGTATCATCAGGAGTAGATTTGAGTCTTATTCGACCAACATCAAGAGAGG 4447
QY 1449 TACAATGGTTATTATTAATCAATGAGGAAACCTTCTTCGAGAGTGTCTAGGCTTCTTTTC 1508
Db 4448 TACAATGGTTATTATTAATCAATGAGGAAACCTTCTTCGAGAGTGTCTAGGCTTCTTTTC 4507
QY 1509 AAACAAGAGATGATCAACACCATTTCTTGAGCAACACGAGAGCTCATGATTAATATC 1568
Db 4508 AAACAAGAGATGATCAACACCATTTCTTGAGCAACACGAGAGCTCATGATTAATATC 4567
QY 1569 GATCATCAAGATGCTGTTTCGGATGATTTCGTTACTGTTTGTGGAATGTTGTGTTAT 1628
Db 4568 GATCATCAAGATGCTGTTTCGGATGATTTCGTTACTGTTTGTGGAATGTTGTGTTAT 4627
QY 1629 GGTGGTTATCAAGATTGTCAGCCCGGTTAACTGCGATGCTTACGCTGCTAGTAGTTT 1688
Db 4628 GGTGGTTATCAAGATTGTCAGCCCGGTTAACTGCGATGCTTACGCTGCTAGTAGTTT 4687
QY 1689 GATTATAACCAAGAAACCAATTTACTGTTGCTCAGCAGCAGACCCAGCAGTCCCA 1748
Db 4688 GATTATAACCAAGAAACCAATTTACTGTTGCTCAGCAGCAGACCCAGCAGTCCCA 4747
QY 1749 GGTGAGATTTCCTCCGCGCAATGACGAATTAATGTTGGCTCTTAATATGTTACCATGGG 1808
Db 4748 GGTGAGATTTCCTCCGCGCAATGACGAATTAATGTTGGCTCTTAATATGTTACCATGGG 4807
QY 1809 GAAGTGGTGAGAGTTGCTCCCAATTTACAGTTTGGACCAATTAAGAAATAG 1868
Db 4808 GAAGTGGTGAGAGTTGCTCCCAATTTACAGTTTGGACCAATTAAGAAATAG 4867
QY 1869 TTAAG 1874
Db 4868 TTAAG 4873

RESULT 6
AF317906
LOCUS
DEFINITION Brassica napus AP2/EREBP transcription factor BABY BOOM1 (BBM1)
4873 bp DNA linear
PLN 10-OCT-2002

gene, complete cds.
AF317906
AF317906.1 GI:21069054
KEYWORDS
SOURCE
ORGANISM
Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Eudicotyledonae;
1 (bases 1 to 4873)
Boutillier,K., Offringa,R., Sharma,V.K., Kieft,H., van
Laumen,A.A.M., Ouellet,T., Zhang,L., Hattori,J., Liu,C.-M.,
Miki,B.L.A., Custers,J.B.M. and van Lookeren Campagne,M.M.
Ecopic expression of BABY BOOM triggers a conversion from
vegetative to embryonic growth
Plant Cell 14 (8), 1737-1749 (2002)
22163009
12172019
2 (bases 1 to 4873)
Boutillier,K. and Fiers,M.
Direct Submission
Submitted (02-NOV-2000) Business Unit Plant Development and
Reproduction, Plant Research International, Droevendaalsesteeg 1,
Wageningen 6708 PB, The Netherlands
Location/Qualifiers
1. 4873
/organism="Brassica napus"
/mol_type="genomic DNA"
/cultivar="Topas"
/db_xref="taxon:3708"
1620..4858
/gene="BBM1"
join(1620..1845,2299..2719,2953..3035,3161..3169,
3315..3403,3554..3627,3798..3848,3962..4038,4149..4858)
/gene="BBM1"
/product="AP2/EREBP transcription factor BABY BOOM1"
join(1620..1845,2299..2719,2953..3035,3161..3169,
3315..3403,3554..3627,3798..3848,3962..4038,4149..4858)
/gene="BBM1"
/codon_start=1
/product="AP2/EREBP transcription factor BABY BOOM1"
protein_id="AA033802.1"
/db_xref="GI:21069055"
/translation="MNNWLGSLSPYEQNHRRKDVYSGSTTTTVDVAGEVCYDPTAA
SDSASATQSPFPFVVVDAFTDNNSHSDWDINGCANINHDNDGPKLENFLG
RTTYTINENVGSGSGCYGGGGGSLGSLMKTWLNRPVNDVNDGNNAAK
GLSLMSNSTCDNNNDNNVAGKTIIDSVATPKTIESFQRTSIYRGVTRHR
WTGVEAHLWNSCKREGQTRGVYLGVDKEEKAARADLAALKYWGTTTTFP
MSBEKEVEEMKEMTRQEVVASLREKSGFSGASLYRGVTRHQRWQARIQVAG
NKDLYLGTFTGQEEAAYDIAAKFRGLTATVTFPMNRVNVKALLESFSLPIGSAK
RLKANRPVPSMMMLSNVSENSASGQNAQVQHFQGVLDLSLHQERNGYYN
CGNLSSEARACFKQEDDQHFHLSMTNIDHQSVDSDSVTCGVVYGYGQ
PAAPVNCDAVAASBFDYNARNHYVFAQQQTQSPGDFPAAMTNNVGSNNMYHGE
GEVAPTFTVNDN"

Query Match 36.0%; Score 726; DB 8; Length 4873;
Best Local Similarity 100.0%; Pred. No. 1.3e-162;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1149 GGCACACAGAAGAGCTGAGAGGATACGACATTCGGCCCATCAAAATTCAGAGGATTA 1208
Db 4148 GGCACACAGAAGAGCTGAGAGGATACGACATTCGGCCCATCAAAATTCAGAGGATTA 4207
QY 1209 ACCGAGTGAATTAATTCAGATGAAACAGATCAACGTTAAAGCAATCTCGAAGCCCT 1268
Db 4208 ACCGAGTGAATTAATTCAGATGAAACAGATCAACGTTAAAGCAATCTCGAAGCCCT 4267
QY 1269 AGTCTTCCTATTGTTAGCGCGCAGAAACGTCCTCAAGAGGCTCAACGTCGCGTTCACAGT 1328
Db 4268 AGTCTTCCTATTGTTAGCGCGCAGAAACGTCCTCAAGAGGCTCAACGTCGCGTTCACAGT 4327
QY 1329 ATGATGATGATCAATTAATTCAGATGAAACAGATCAACGTTAAAGCAATCTCGAAGCCCT 1388


```

/gene="T10B6_30"
/number=3
complement(6118..6197)
/gene="T10B6_30"
/number=3
complement(6198..6255)
/gene="T10B6_30"
/number=4
complement(6256..6341)
/gene="T10B6_30"
/number=4
complement(6342..6432)
/gene="T10B6_30"
/number=5
complement(6433..6605)
/gene="T10B6_30"
/number=5
complement(6606..6672)
/gene="T10B6_30"
/number=6
complement(6673..6860)
/gene="T10B6_30"
/number=6
complement(6861..6961)
/gene="T10B6_30"
/number=7
complement(6962..7117)
/gene="T10B6_30"
/number=7
complement(7118..7220)
/gene="T10B6_30"
/number=8
complement(7221..7368)
/gene="T10B6_30"
/number=8
complement(7369..7579)
/gene="T10B6_30"
/number=9
complement(7580..7733)
/gene="T10B6_30"
/number=9
complement(7734..7864)
/gene="T10B6_30"
/number=10
complement(7865..7948)
/gene="T10B6_30"
/number=10
complement(7949..8092)
/gene="T10B6_30"
/number=11
complement(8093..8232)
/gene="T10B6_30"
/number=11
complement(8233..8357)
/gene="T10B6_30"
/number=12
complement(8358..8456)
/gene="T10B6_30"
/number=12
complement(8457..8523)
/gene="T10B6_30"
/number=13
complement(8524..8724)
/gene="T10B6_40"
complement(join(8724..9038,9121..10524))
/gene="T10B6_40"
complement(join(8724..9038,9121..10524))
/gene="T10B6_40"
/note="similarity to 2-hydroxyphytanoyl-CoA lyase - Homo sapiens, EMBL:AJ31753 Contains Prokaryotic membrane lipoprotein lipid attachment site AA100-110;Thiamine pyrophosphate enzymes signature AA440-459"

```

```

/codon_start=1
/product="2-hydroxyphytanoyl-CoA lyase-like protein"
/protein_id="CAC01733.1"
/db_xref="GI:9755761"
/db_xref="GOA:Q9LF46"
/db_xref="SPTREMBL:Q9LF46"
/translation="MADKSETTPPSIDGNVLVAKSLHSLGVTHMFGVGVIPVTSLASR
AMALGIRFIAFHNEQASAGYAASGYLTGKPGILLTVSPGCVHGLAGLSNAWNTWP
WMTSGCQORDVGRGDFQELQIEAKAFSLSKAKDVREI PCVSRVLDRVSGR
PGCYLDIDTDLVRQKISSEADKLDEVERSKKEP RGSRLSEIERSLAKGR
PLVFGKRAASADELKLVEITIGFPIPMGKGLLFDTHESATARSALAKKC
DVALVGARLNLWLLHGESPKDKVKFILDVSEEEIELRPHLIGVGDAKTVIGLL
NREIKDDPFLGKSNWVESISKAKENKMEIQAKDVFFELTMRIRDAILA
VEGSPVWSEGAANTMDVGRVSLVQKEPTLDAGTGMVGLGYCTAAASVADPRL
WVAVGSDSGFGFSAEMVELVRYNLAVIIVFNNGCVGDRRGEESGPHKEDPAP
TSFVPNAGYHKLIEAFGGGYIVETDDELKSALAESFAARKPAVVNVIIDPFAGESG
RLOHKN"
complement(8724..9038)
/gene="T10B6_40"
/number=1
complement(9039..9120)
/gene="T10B6_40"
/number=1

```

Query Match	Best Local Similarity	Score	504.8;	DB 8;	Length	33563;			
Matches	703;	Conservative	0;	Mismatches	137;	Indels	46;	Gaps	7;
QY 1145	TTTTGGCACACAAAGAGAGCTCGAGGAGCATACGATTCGGGCGCATCAAAATTCAGAGG	1204							
DB 27075	TGTAGGCACACAGAGAGGCTGCTAGGCTTATGATCAATTCAGCCATTAATTCAGAGG	27016							
QY 1205	ATTAACCGCAGTGACTTAACCTTCGACATGACACATACACGTTAAAGCAATCTCGAAG	1264							
DB 27015	ATTAGCGCAGTGACTTAACCTTCGACATGACACATACACGTTAAAGCAATCTCGAAG	26956							
QY 1265	CCCTAGTCTTCTTATGTTAGTGGCCGCAAAACCTCTCAAGAGGCTTAAACGTCGGTTC	1324							
DB 26955	CCCGAGTCTACCTATTGGTAGTTCTCGCAAACTCTCAAGGACGTTAATAATCCGGTTC	26895							
QY 1325	AAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1384							
DB 26895	AGCT---ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	26839							
QY 1385	AAACGTCGGGTTGAGCATCATCAGGAGTAGATTTGAGCTTATTCGACCAACATCAAGA	1444							
DB 26838	AAACACTCGGTTTCAGCATCATCAGGAGTAGATTTGAGCTTATTCGACCAACATCAAGA	26779							
QY 1445	GAGGTACAATGGTTATTATTACAATGAGGAAACTTGTCTTCGAGAGTGTCTAGGCTTG	1504							
DB 26778	GAGGTAGTTGG---TTATTACAATGAGGAAACTTGTCTTCGAGAGTGTCTAGGCTTG	26722							
QY 1505	TTTCAAAACAAGAGGATGATCAACACCATTTCTTTGAGCAACAACGAGGCTCATGCTAA	1564							
DB 26721	TTTCAAAACAAGAGGAGGAGCAACACACTTCTTTGAGAAACTCCCGGAGTCACATGCTAA	26662							
QY 1565	TATCGATCATCAAGTTCTGTTTCGGATGATTCGGTTACTGTTTGGAAATGTTCTTGG	1624							
DB 26661	TGTTGATCATCATGCTCGACCTCTGATGATTTGTTCCGTTTGGAAATGTTCTTGG	26602							
QY 1625	TTATGGTGGTTATCAAGGATTTGAGGCC-----CGTTAACTGGGATGCCTTA	1672							
DB 26601	TTATGGTGGTTATCAAGGATTTGAGGCC-----CGTTAACTGGGATGCCTTA	26542							
QY 1673	CGCTGCTAGTGAGTTGATTAACGCAAGAAACCATTTACTTTGCTCAG-----	1724							
DB 26541	TACTGCTGCTGAGATTCCTTACACCGCAAGAAATCATTTACTATGCTCAGCATCAGCA	26482							
QY 1725	-CAGCAGACCCAGCAGTCGCGAGTGAGATTTTCGCGGCGCAATCAGCAATATGTT	1783							
DB 26481	ACAACAGCAGATTCAGCAGTCGCGGAGGAGATTTTCGCGGCGCAATTCGAATAACCA	26422							
QY 1784	TGGCTCTAATATGATTATACCATGGGAGAGGTGGTGAGAGAGTTGCTCCCAACATTTACAGT	1843							

Db	4780	TTATGGTGGTTATCAGAGTTCGCAATCCCTGTTGGAAACATCGTGTAAATTACGATCCCTT	4831
Qy	1673	CCCTCTAGTGGTTGATTATACGCAAGAAACCATATTACTTTCTCAG-----	1724
Db	4840	TACTGCTGCTGAGATTGCTTACAAACGCAAGAAATCATTTACTATGCTCAGCATCAGCA	4899
Qy	1725	-CAGCAGCAGACCCAGCAGTCGCCAGGTGGAGATTTTCCCGCGCAATGACGAATATGT	1783
Db	4900	ACAAACAGCAGATTACAGAGTCGCCGGGAGGAGATTTCGGTGGCGATTTCGATTAACCA	4959
Qy	1784	TGGCTCTAATATGATTATACCATGCGGGAAGGTGGTGGAGAAAGTTGCTCCACATTTACAGT	1843
Db	4960	TAGCTCTTAACTATGATCTTACCGGGGAAGGTGGTGGAGAAAGGGCTCCAAACGTTTTCAGT	5019
Qy	1844	TTGGAAACGACAATTAGAAAAATAGTTAAAGATCTTTTAGTTATATGATGGTGTGTTGTGCT	1903
Db	5020	TTGGAAACGACACTTAGAAAAATAGTAAAGATCTT---TTAGTTGTTGCTTTGATGT	5076
Qy	1904	GGTGAACAGTGTGATCATCTTGAATTATGTTTTTTTCTTCTCTTTCTTTCTTTCTTGTTA	1963
Db	5077	TGCGAACAGTTGATCTCTGTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	5136
Qy	1964	ATT 1966	
Db	5137	TTT 5139	
RESULT 10			
AX048247			
LOCUS	AX048247	1749 bp	DNA linear PAT 15-DEC-2000
DEFINITION	Sequence 3 from Patent WO0070058.		
ACCESSION	AX048247		
VERSION	AX048247.1	GI:11877017	
KEYWORDS	Arabidopsis sp.		
SOURCE	Arabidopsis sp.		
ORGANISM	Arabidopsis sp.		
REFERENCE 1	Turner, S. and Taylor, N.		
AUTHORS	Turner, S. and Taylor, N.		
TITLE	Plant cellulose synthase genes		
JOURNAL	Patent: WO 0070058-A 3 23-NOV-2000;		
FEATURES	THE VICTORIA UNIVERSITY OF MANCHESTER (GB)		
source	Location/Qualifiers		
ORIGIN	1..1749		
Query Match	19.8%;	Score 399.2;	DB 6; Length 1749;
Best Local Similarity	77.8%;	Pred. No. 1.9e-84;	
Matches	602;	Conservative 0;	Mismatches 123; Indels 49; Gaps 8;
Qy	1257	CTCGAAAGCCCTAGTCTTCCTATTGTTAGTCCGCCCAAAAGCTCTCAAGGAGGCTTAAACCGT	1316
Db	1	CTCGAGAGCCCGAGTCTACCTATTGTTAGTTCTCGGAAACGCTCTCAAGGAGGTTAAACAAT	60
Qy	1317	CCGGTTCAGATGATGATCATGATCAGTATACGTTTTCAGAGAGTGAATAGTGTAGC	1376
Db	61	CCGGTTCAGCT---ATGATGATTAGTAAATACGTTTTCAGAGAGTGAATATGTTAGC	117
Qy	1377	GGTTGGCAAAAACCGTCCGCTTTCAGCATCATCAGGAGGTAGATTTGAGCTATTTCACCAA	1436
Db	118	GGTTGGCAAAAACCTGCGCTTTTCAG---CATCAGGAATGGATTGAGCTATTTCAGCAA	174
Qy	1437	CATCAGAGAGGTACATGTTATTATATACATGAGGAACCTTGTCTTCGGAGAGTGT	1496
Db	175	CAGCAGGAGAGTACGTTGG---TTATTACAAATGAGGAACCTTGTCTTCGGAGAGTACT	231
Qy	1497	AGGGCTGTCTTTCACAAACAGAGGATGATCAACACCATTTCTTTGAGCAACACCGCAGAGCTC	1556

Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones

TITLE JOURNAL REFERENCE AUTHORS

2 (bases 1 to 1699)

Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission

Submitted (03-JUN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES source

1..1699
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosomes="4"
/clone="U11873"
/ecotype="Columbia"
/notes="This clone is in pUNI 51."
1..1699
/gene="At4g37750"
1..1668
/gene="At4g37750"
/codon_start=1
/evidence="experimental"
/product="putative ovule development protein aintegumenta"
/protein_id="AA051282.1"
/db_xref="GI:21436287"
/translation="MKSFDNDNNHNTNLLGFLSNMNMKGGRGREAIYSSST
SRAATSSVPVPLVGDNTNFGVCGYGNPNNGGIYSHNVMPLSDSLCMEALNR
SSHNHDSQSPKVEDPFGTHNNTSHKAMDLSLFLYNTHPNTTNQEFESP
PQTRHEETRNYPDPSLTHGGSFNVGVYFQSSLSMSPGSSCITGSHHQH
NQNRQSQNHQOISALVETSVGFETTTMAAKKRGQEDVYVVGQKQVHRKSIDT
FGQTSOYRGVTRHRTGYEAHLWDNSFFKRGHSGRQVLYGGVDEEKAARAYDL
AALKYWGSPHTNFSAENYQKETEDMKNTRQBYVAHLRKSGSFGSRGASIVRGVTRH
HQHGRQARCVAGNKDLYLGFQTEEAADYVAALKFRGTNAVTFDI TRYDVD
RIMSSNTLSGELARNNNSIVVRNEDQTLNNAVVEGSGNKEVSPPELLLSFPAIFA
LPQVNMFGSNGMGNPWTSPNPAELKTVALTLPQMPVFAWADS"
1669..1699
/gene="At4g37750"

gene CDS

16.48; Score 330.6; DB 8; Length 1699;
Best Local Similarity 75.4%; Pred. No. 4.9e-68;
Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

3'UTR ORIGIN

Query Match 16.48; Score 330.6; DB 8; Length 1699;
Best Local Similarity 75.4%; Pred. No. 4.9e-68;
Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 714 AAAAATATTGAGAGTTTGGACAGAGAGCGTCTATATACCGCGGTGTTTCAAGGCGATCGG 773
Db 814 AATATCTATCGATCTTTTGGACACAGCACTTCTCATATCCGAGCGGTTCACAGCATAGA 873
QY 774 TGGACAGAGATATAGGCGACATTTATGGGATATAGTTGTTAAAGAGAGAGGCCAAAGC 833
Db 874 TGGACTGTTAGATATGAAGCTCATCTATGGGACAAATAGTTTCAAGAGGAGGTACAGT 933
QY 834 CGCAAGAGAGAGCAAGTTTATTTTGGGAGGTTATGACAAAGAGAGAGAGAGAGAGGCT 893
Db 934 AGAAGAGAGAGAGAGTTTATCTGGGAGGTTATGATATGGAGGAGAGAGAGTGTCTCGAGCA 993
QY 894 TATGATTATCCCGCACTCAAGTATTTGGGAAACCAACCACTACTACTACTTCCCATGAGC 953
Db 994 TATGATCTTCTGCACTCAAGTACTGGGTCCCTACTACTCACCAATTTCTCTGCGGAG 1053
QY 954 GAATATCAAAAAGAGGTAGAGAGATGAACACATGACAGGCAAGAGTAGTATGTTGCTCA 1013
Db 1054 AATATCAGAAAGAGATGAAGACATGAAGACATGACTAGACAGATATGTTGCACAT 1113
QY 1014 CTGCGCAGGAAAAGTAGTGTGTTTCTCTCGTGGTGCATCGATTTTATCGTGGAGTAACAAGA 1073
Db 1114 TTGAGAGAGAGAGAGAGTGTGTTTCTCTAGGGTGTCTTCCATCTATAGAGGAGTACACAAGA 1173
QY 1074 CATCACCAACATGGAAGATGCGAGCTAGCATAGGAGAGTCCCGGTACACAGACCTC 1133
Db 1174 CATCACAGCATGGAAGGTGGCAAGCAGCGATTGGTGTAGTTCGCTGAAAACAAGATCTC 1233
QY 1134 TACTTTGGGAACCTTTTGGCACACAGAAAGAGCTGCAGAGGCATACGACATTTGGGCGCATC 1193
Db 1234 TACCTTGGNACTTTTGGACCCAGAGAGAGTTCAGCAAGCTTACGATGTAGCAGCAATT 1293
QY 1194 AAATTCAGAGATTAAACCGAGTGCATTAAGTTCAGATGACATGACAGATACACAGTTAAAGCA 1253
Db 1294 AAGTTCCGTGGCACAAATGCTGTGACTTAACCTTTGATATCAAGAGTACGATGTTGATCGT 1353
QY 1254 ATCCT 1258
Db 1354 ATCAT 1358

RESULT 13

ATU41339
LOCUS Arabidopsis thaliana ANT (AINTEGUMENTA) mRNA, complete cds.
DEFINITION Arabidopsis thaliana ANT (AINTEGUMENTA) mRNA, complete cds.
ACCESSION U41339
VERSION U41339.1 GI:1244707
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE AUTHORS

Elliot, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q., Gerentes, D., Perez, P. and Smyth, D.R.

AINTEGUMENTA, an APTALA2-like gene of Arabidopsis with pleiotropic

roles in ovule development and floral organ growth

Plant Cell 8 (2), 155-168 (1996)

JOURNAL 96351414

MEDLINE 8742707

PUBMED 2 (bases 1 to 1905)

REFERENCE Smyth, D.R.

AUTHORS Direct Submission

TITLE Submitted (27-NOV-1995) David Smyth, Genetics and Dev. Biology,

JOURNAL Monash University, Wellington Road, Clayton, VIC 3168, Australia

FEATURES Location/Qualifiers

1..1905

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strains="Landsberg erecta"

/db_xref="taxon:3702"

```

/chromosome="4"
/map="between spt and cer-9"
/tissue type="flower; ovule"
1..1905
/gene="AINTEGUMENTA"
38..1705
/gene="AINTEGUMENTA"
/note="similar to APETALA2 protein encoded by GenBank
Accession Number U12546"
/codon_start=1
/product="ANT"
/protein_id="AABI7364.1"
/db_xref="GI:1244708"
/translation="MKSFCDNDNNHNTNLLGFLSSNMKMGGRGREATYSSST
SSAATSSSVPPQVVGDNSTNFVCYCSNPNGGIYSHMVMPLRSDGSLCLMEALNR
SSHNHHDQSSPKVEDPFGTHNNNTSHKAMDLSLDFYNTTHEPNTTNFQFTSF
POTNHEETENYKNDPSLTHGGSFNVGVGEFOOSLSLSPGSSOCCITGSHHQO
NONHONSONHQOISEALVETSVGFETTAAAKKKGQEDVVVGKQIVHRKSIDT
FGQTSQYGVTRHWTRGYEAHLWDSNFKKGRSGRQVYLGGYDMEKAAAYDL
AALKWGPSTHNSAENYKQIEDMKNTMQEYVAHLRRKSSGSRGASIIYGVTRH
HQRGWARIGRVAGNKDLYLGTQTEEAAYDVAAIKFRGTNAVTFDITRYDVD
RIMSNTLLSGELARNNSIVVRNTEDQTLNNAVVEGSKNEVSTPERLLSPFAIFA
LPQVNMFGSNMGMNMPSTNPNELKTVALTLPMPVFAAWDS"
878..1108
/gene="AINTEGUMENTA"
/note="encodes first AP2 domain"
1184..1391
/gene="AINTEGUMENTA"
/note="encodes second AP2 domain"

misc_feature
16.4%; Score 330.6; DB 8; Length 1905;
Best Local Similarity 75.4%; Pred. No. 4.9e-68;
Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 714 AAAAATAATGAGAGTTTGGACAGACATGACGCTATATACCGCGGTGTACAGGCATCGG 773
DB 851 AAATCTATCGATCTTTTGGACAGACATCTCAATACCGAGCGTTTCAAGACATAGA 910
QY 774 TGGACAGACATATGAGGACATTTTGGATAATAGTTGTAAGAGAGAGCCAAACG 833
DB 911 TGGACTGATAGATGAAGTCACTATGAGCAATAGTTTCAAGAGAGGAGTCAAGT 970
QY 834 CGCAAGGAGACAAAGTTTATTTGGAGGTTTATGACAAAGACAAAGACAGTACGGCT 893
DB 971 AGAAAGGAGACAAAGTTTATCTGGAGGTTATGATATGAGGAGAGAGCTGTCGAGCA 1030
QY 894 TATGATTTAGCGGACTCAAGTATTTGGGAAACCACTACTACTACTACTACTACTACTAGC 953
DB 1031 TATGATCTTCTGCACTCAAGTACTGGGTCCCTCTACTCACCACCAATTTCTCTGGGAG 1090
QY 954 GAATATGAAAGAGGTAGAGAGATGAAGACATGACAGGACGAGATGTTGCTCTCA 1013
DB 1091 AATTATCAGAGAGATTTGAGACATGAAGAAATGACTAGACAGAAATATGTCACAT 1150
QY 1014 CTGGCAGGAAAGTATGTTGTTCTCTCGGTGCAATGATTTATCGTGAGTAAACAGA 1073
DB 1151 TTGAGAGGAGAGCAGTGGTTTCTCTAGGGGTGCTTCCATCTATAGAGAGTCAAGA 1210
QY 1074 CATCACCAATGGAAGTGGCAAGTATGATAGGAGAGTCCCGGTAAACAGACCTC 1133
DB 1211 CATCACCAATGGAAGTGGCAAGCAGGATTTGATAGTCTGTAACAAAGATCTC 1270
QY 1134 TACTTGGGAACTTTTGGCAACAAGAGAGTTCGAGAGGCATACGACATTCGGCCATC 1193
DB 1271 TACCTTGGAACTTTTGGAAACCAAGAGAGTTCGAGAGCTTACCATGAGCAAT 1330
QY 1194 AATTTCAGAGGATTAACCGAGTACTTCTGATCTGATGATGATGATGATGATGATGATGAT 1253
DB 1331 AGTTTCGTGGCAAAATGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1390
QY 1254 ATCCT 1258
|||||

```

```

Db 1391 ATCAT 1395

RESULT 14
ATU44028
LOCUS
DEFINITION
Arabidopsis thaliana transcription factor CKC mRNA, complete cds.
ACCESSION
U44028
VERSION
U44028.1 GI:1171428
KEYWORDS
SOURCE
Arabidopsis thaliana (chale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1955)
Vergani, P., Morandini, P. and Soave, C.
Direct Submission
Submitted (29-DEC-1995) Paola Vergani, Dept. of Biology 'L.
Gorini', University of Milan, Via Celoria 26, Milan, 20133, Italy
Location/Qualifiers
1..1955
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
76..1743
/function="transcription factor; complements a yeast
protein kinase C 1 mutant; Method: conceptual translation
supplied by author."
/codon_start=1
/product="CKC"
/protein_id="AAA86281.1"
/db_xref="GI:1171429"
/translation="MKSFCDNDNNHNTNLLGFLSSNMKMGGRGREATYSSST
CSAATSSSVPPQVVGDNSTNFVCYCSNPNGGIYSHMVMPLRSDGSLCLMEALNR
SSHNHHDQSSPKVEDPFGTHNNNTSHKAMDLSLDFYNTTHEPNTTNFQFTSF
POTNHEETENYKNDPSLTHGGSFNVGVGEFOOSLSLSPGSSOCCITGSHHQO
NONHONSONHQOISEALVETSVGFETTAAAKKKGQEDVVVGKQIVHRKSIDT
FGQTSQYGVTRHWTRGYEAHLWDSNFKKGRSGRQVYLGGYDMEKAAAYDL
AALKWGPSTHNSAENYKQIEDMKNTMQEYVAHLRRKSSGSRGASIIYGVTRH
HQRGWARIGRVAGNKDLYLGTQTEEAAYDVAAIKFRGTNAVTFDITRYDVD
RIMSNTLLSGELARNNSIVVRNTEDQTLNNAVVEGSKNEVSTPERLLSPFAIFA
LPQVNMFGSNMGMNMPSTNPNELKTVALTLPMPVFAAWDS"

ORIGIN
Query Match 16.4%; Score 330.6; DB 8; Length 1955;
Best Local Similarity 75.4%; Pred. No. 4.9e-68;
Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 714 AAAAATAATGAGAGTTTGGACAGACATGACGCTATATACCGCGGTGTACAGGCATCGG 773
DB 889 AAATCTATCGATCTTTTGGACAGACATCTCAATACCGAGCGTTTCAAGACATAGA 948
QY 774 TGGACAGACATATGAGGACATTTTGGATAATAGTTGTAAGAGAGAGCCAAACG 833
DB 949 TGGACTGATAGATGAAGTCACTATGAGCAATAGTTTCAAGAGAGGAGTCAAGT 1008
QY 834 CGCAAGGAGACAAAGTTTATTTGGAGGTTTATGACAAAGACAAAGACAGTACGGCT 893
DB 1009 AGAAAGGAGACAAAGTTTATCTGGAGGTTATGATAGGAGAGAAAGCTGTCGAGCA 1068
QY 894 TATGATTTAGCGGACTCAAGTATTTGGGAAACCACTACTACTACTACTACTACTACTAGC 953
DB 1069 TATGATCTTCTGCACTCAAGTACTGGGTCCCTCTACTCACCACCAATTTCTCTGGGAG 1128
QY 954 GAATATGAAAGAGGTAGAGAGATGAAGACATGACAGGACGAGATGTTGCTCTCA 1013
DB 1129 AATTATCAGAGAGATTTGAGACATGAAGAAATGACTAGACAGAAATATGTCACAT 1188
QY 1014 CTGGCAGGAAAGTATGTTGTTCTCTCGGTGCAATGATTTATCGTGAGTAAACAGA 1073
DB 1189 TTGAGAGGAGAGCAGTGGTTTCTCTAGGGGTGCTTCCATCTATAGAGGAGTCAAGA 1248

```

QY	1074	CATCACCAATGGAAGATGGCAAGCTAGATAGGAGAGTCCCGGTACCAAGACCTC	1133
Db	1249	CATCACCAATGGAAGATGGCAAGCTAGATAGGAGAGTCCCGGTACCAAGACCTC	1308
QY	1134	TACTTGGGAATTTTGGCACACAAGAGAGCTGCAGAGGCATACGACATTCGGCGCATC	1193
Db	1309	TACTTGGGAATTTTGGAACCCAGAGAGCTGCAGAGCTTACGATGTAGCAGCAATT	1368
QY	1194	AAATTGAGAGGATTACCGCAGTACATTAATTCGACATGACATACAGATACAGCTTAAGCA	1253
Db	1369	AAATTGCGTGGCAAAATGCTGTGACTAACTTTGATATCATCAGAGTACGATGTGTCGT	1428
QY	1254	ATCCT 1258	
Db	1429	ATCAT 1433	
RESULT 15			
AY080706			
LOCUS			
DEFINITION	AY080706 2056 bp mRNA linear PLN 18-SEP-2002		
ABRIDGED	Arabidopsis thaliana putative ovule development protein		
ACCESSION	AY080706		
VERSION	AY080706.1		
KEYWORDS	FUT CDNA		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 2056)		
AUTHORS	Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		
TITLE	Arabidopsis Full Length cDNA Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2056)		
AUTHORS	Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL CDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
The Salk, Stanford, PGE (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.			
Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs.			
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.			

FEATURES		Location/Qualifiers	
source		1..2056	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/chromosome="4"	
		/clone="RAPL06-86-F19 (R11873)"	
		/ecotype="Columbia"	
		/note="This clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SstI insert."	
gene		1..2056	
		/gene="At4g37750"	
5' UTR		1..157	
		/gene="At4g37750"	
misc_difference		2	
		/gene="At4g37750"	
		/note="not present in genomic sequence"	
CDS		158..1825	
		/gene="At4g37750"	
		/codon_start=1	
		/evidence=experimental	
		/product="putative ovule development protein aintegumenta"	
		/protein_id="AAL85024.1"	
		/db_xref="GI:19310587"	
		/translation="MKSFCDNDNNHNTNLLGFLSSNNMKMGGRGGRALVSSSTSSAATSSVPPQVVGDNSTNFGVCGSNPGVGHVSHVMPRLSDGLCLMELNRSNHNHSDSPKQVDFGTHNNTSHKEMDLSDLFYNTTPEPTNTNFQFFHSFPTNRHETERNVGNPSLTHGGSFNVVGFQSLSLMSPGSQSCQIIVRSHQNONHOSQHQIIEALVETSGFPTTAAAKKKGQEDVVVVGKQIIVRKSIDTFCQRTSQVGRVTRHRTGRVREAHLDNSPKKESHRKGRQVLYGGYNEEKAARAYDLAALVGFSTHTNFSAENYOKIEIDNNNNQEVYVHLRRKSGFSKASLYRGVTRHQAQWRQARVAGNKDLYLGTQTEBAEAYDVAALKFRGNVNTFDITSDYDRIMSNTLLSGELARRNNSTIVRNTDQALNAVVEGSGNKEVSTTPEPTNTNFQFFHSFPAIFALPQVQKMFSGMGNMSPWTSNENALTKVTLTPQMPVFAAWDS"	
		1826..2056	
		/gene="At4g37750"	
3' UTR			
ORIGIN			
Query Match		16.4%; Score 330.6; DB 8; Length 2056;	
Best Local Similarity		75.4%; Pred. No. 4.9e-68;	
Matches 411; Conservative		0; Mismatches 134; Indels 0; Gaps 0;	
QY	714	AAACTATTGAGAGTTTGGACAGAGCGTCTATATACCGCGTGTACAGGCATCGG	773
Db	971	AAATCTATCGATATCTTTGGACACAGCACTTCTCAATACCGCGGTGTACAGGCATAGA	1030
QY	774	TGGACAGCAAGATATGAGGCACATTTATGGGATATATAGTTGTAAGAGAGGCGCAACG	833
Db	1031	TGGACTGTGATATATGAGCTCATCTATGGGACATATAGTTTCAAGAGGAGGTACAGT	1090
QY	834	CGCAAGAGCAACAGTTTATTTGGAGGTTATGACAAAGAGAGAGAGAGAGCTAGGCT	893
Db	1091	AGAAAGAGAGCAAGTTTATCTGGAGGTTATGATATGGAGGAGAGAGCTGCTCGAGCA	1150
QY	894	TATGATTTAGCGGCACCTCAAGTATTGGGAAACCACTACTACTACTCTCCCATGAGC	953
Db	1151	TATGATCTGTGCACTCAAGTACTGGGTCCTCTACTCACCAATTTCTCTCGGAG	1210
QY	954	GAATATGAAAAGAGGTAGAGAGATGACATGACAGGCAAGAGATATGTTGCCTCA	1013
Db	1211	AATTATGAAAAGAGATTTGAAGACATGAAGAACATGACTAGACAGAAATATGTTGCACAT	1270
QY	1014	CTGGCAGGAAAAGTAGTGTTTCTCTGCTGCTCATCTGATTTATCGTGGAGTAAACAAGA	1073
Db	1271	TTGAGAGAGAGACAGTGTTTCTTAGGGGTGCTTCCATCTATAGAGAGTCAACAAGA	1330
QY	1074	CATCACCAATGGAAGATGCAAGCTAGGATAGAGAGTCGCCGTGTAACAAGAGCCTC	1133
Db	1331	CATCACCAATGGAAGATGCAAGCTAGGATAGAGAGTCGCCGTGTAACAAGAGCCTC	1390
QY	1134	TACTTGGGAATTTTGGCACACAAAGAGAGCTGCAGAGGCATACACATTTGGCGGCATC	1193
Db	1391	TACTTGGGAATTTTGGCACACAAAGAGAGCTGCAGAGGCATACACATTTGGCGGCATC	1450

Oy	1194	AAATTCAGAGGATTAAACCGCAGTGA	AACTTCGACATGAACAGATACAACGTTAAAGCA	1253
Db	1451	AAGTTCCTGGCACAATAATGCTGTGACTAAC	TTTGATATCAGAGGTACGATGTTGATCGT	1510
Oy	1254	ATCCT	1258	
Db	1511	ATCAT	1515	

Search completed: July 19, 2004, 08:51:46
Job time : 7912 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 05:17:49 ; Search time 5287 Seconds
(without alignments)
11375.531 Million cell updates/sec

Title: US-09-980-364-1

Perfect score: 2014

Sequence: 1 gttctctctctctcttaag.....gttgataaatttcagact 20:4

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_esthc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	534.4	26.5	556	B2512981	B2512981 BOMQP05TF
c 2	393.4	19.5	781	BH443187	BH443187 BOGVV76TR
3	382.6	19.0	769	CA783156	CA783156 sac20d05.
4	370.2	18.4	720	BQ864461	BQ864461 QGC26M12.

5	361.4	17.9	590	12	BM094116
6	351.4	17.4	748	14	CF807326
7	346.4	17.2	674	10	BF647766
8	344.4	17.1	552	13	BU964897
9	339.4	16.9	644	10	AW200688
10	339.2	16.8	882	14	CK367021
11	335.4	16.7	500	10	AW780688
12	335.2	16.6	593	12	BM307997
13	333.6	16.6	717	28	BZ044547
14	332	16.5	686	28	BZ496117
15	323.6	16.1	545	14	CD475882
16	317.2	15.7	585	12	BM086088
17	308.6	15.3	631	13	CA103041
18	307.4	15.3	640	13	CA100375
19	307.2	15.3	697	14	CA189006
c 20	306	15.2	481	28	BH694176
21	305	15.1	555	12	BU188928
22	305	15.1	558	12	BU178045
23	304.2	15.1	690	13	CA094556
24	302	15.0	641	14	CD879292
25	300.4	14.9	518	28	BH443196
26	299.8	14.9	1753	11	AY109146
27	297.2	14.8	663	28	BH927892
28	296.6	14.7	767	28	BZ077316
29	288.4	14.3	466	13	BU965263
c 30	273	13.6	417	28	BH908835
31	272.6	13.5	599	14	CA230632
32	271.2	13.5	1160	14	CK206573
33	270.8	13.4	558	14	CA232734
34	268.2	13.3	476	12	BM092890
35	268.2	13.3	545	12	BI974354
36	265	13.2	527	12	BM309412
37	256.4	12.7	432	12	BG882497
38	254.4	12.6	596	12	BM307877
39	253.4	12.6	643	28	BZ516982
40	251.4	12.5	513	13	BQ122372
c 41	251.2	12.5	497	14	CF244784
42	248.4	12.3	842	14	CF446573
43	245.6	12.2	704	13	BQ625052
44	244.6	12.1	469	14	CF324329
45	242.4	12.0	420	9	AJ475492

ALIGNMENTS

RESULT 1
BZ512981/c BZ512981 556 bp DNA linear GSS 16-DEC-2002
LOCUS BOMQP05TF BO_2_3_KB Brassica oleracea genomic clone BOMQP05,
DEFINITION genomic survey sequence.
ACCESSION BZ512981
VERSION BZ512981.1 GI:27039560
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
1 (bases 1 to 556)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOMQP05TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: Tf
Class: sheared ends.

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE AUTHORS

1 (bases 1 to 769)
 Khanna.A., Bella.B., Keim.P., Vodkin.L., Expelding,J., Corvett,V.,
 Shoemaker.R., Marra.M., Hillier.L., Kucaba.T., Martin.J.,
 Beck.C., Wylie.I., Underwood.K., Steptoe.M., Theising.B., Allen.M.,
 Bowers.Y., Person.B., Swaller.I., Gibbons.M., Pape.D., Harvey.N.,
 Schurk.R., Ritter.E., Kohn.S., Shin.T., Jackson.Y., Cardenas.M.,
 McCann.R., Waterston.R. and Wilson.R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 454.

FEATURES source

1..769
 Location/Qualifiers
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl036-14361"
 /tissue_type="somatic embryos cultured on MSD 20"
 /lab_host="DH10B"
 /clone_lib="Gm-cl036"
 /note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This
 cDNA library was constructed from mRNA isolated from
 somatic embryos (age ranging from 2 months to 9 months)
 cultured on MSD 20. The library was prepared using the
 Life Technologies pSuperScript cDNA library construction
 kit. Complementary DNA was synthesized from mRNA using a
 poly (dT) sequence with a NotI restriction site. SalI
 linkers adapters were ligated to the blunt-ended cDNA
 fragments followed by NotI digestion. The cDNA fragments
 were directionally cloned into the NotI-SalI restriction
 site of the pSPORT1 vector. The ligated cDNA fragments
 were transformed into E.coli Electromax DH10B host cells.
 This library was constructed in the laboratory of Dr. Lila
 Vodkin by Anu Khanna at the University of Illinois at
 Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Query Match 19.0%; Score 382.6; DB 14; Length 769;
 Best Local Similarity 76.6%; Pred. No. 2.le-66;
 Matches 482; Conservative 0; Mismatches 144; Indels 3; Gaps 1;
 692 TAGCGTTGAAGCTACACCGAGAAACTATTGAGCTTTGGACAGACGCTCATATA 751
 94 TACTGTTGAATAGCACTAGAGAACTTTGGATACATTCGGCGAGAACATCCATATA 153
 752 CCGGGGTGTTTACAAGGATCCGTTGGACGAGGAATATGAGGCACATTTATGGGATAATAG 811
 154 TCGTGGAGTAACTCGACATAGATGAGCTGGAAGGTATGAAGTCATCTTTGGGATAATAG 213
 812 TTGTAAAGAGAGCGCCAAACCGGCAAGAGAACAGAGTTTATTTGGAGGTTATGACAA 871
 214 CTGTAGAAGGAGGCGCAATCAAGAAAGGACGCCAAGTTTATTTGGGTGGATATGATAA 273
 872 AGAAGAAAAGCAGCTAGGCTTATGATTTAGCGGCACTCAAGTATTTGGGAAACCAACAC 931
 274 AGAAGAGAAGCAGCTAGACCTTATGATTTAGTCTCTGAATCTCTGGGGACATCCAC 333
 932 TACTACTAATCTCCCATGAGCGGAATATGAAAAGAGGTTAGAGAGATGAGACATGAC 991
 334 CACTTACCACCTTCTCCATTTAGCAACTATGAGAGGAATTTGGATGAAATGAACACATGAC 393

992 AGGCGAAGAGTATGTTCTCTCACTGGCGAGAAAGTAGTGTGTTCTCTCGTGGTGCATC 1051
 394 GAGACAAGAAATTTGTTGCCGCCATTTAGAGGAAAGACAGTGGTTTCTCCAGGGGTGCAIC 453
 1052 GATTATTCGTGGAGTAAACAAGACATCACCAACATGGAAGATGCAAGCTAGGATAGGAAG 1111
 454 AATGTATCGTGGAGTTTACAAGGCATCACCAACACGGAAGATGCAAGCAAGGATTGGCAG 513
 1112 ATGCGCGGTAAACAAGACCTCTTCTTGGAACTTTTGGCACACACAGAGAGAGCTGCAGA 1171
 514 AGTTGCGAGAAACAAGATCTTTTACTTGGGAATTTTCACTAGTACGAGAGAGGCTGCAGA 573
 1172 GGCATACGACATTTGGCGCCATCAAAATTCAGAGGATTAACCGCAGTCACTAACTTCGACAT 1231
 574 AGATACGACATAGACGCGATAAAGTTTCAGAGGTCTCAACGCTGTCAACAATTTTCACAT 633
 1232 GAACAGATACACGTTTAAAGCAATCTTCGAAAGCCCTAGTCTTCCTAT---TGTGAGCGC 1288
 634 GAGCGGATACGAGTGAAGCCATTTCTGAAAGCAACACTCTCCCAATAGGAGGAGCGCG 693
 1289 CGCAAAACGCTCAAGGAGGCTAACCGTC 1317
 694 TGAACGCGTCTGAAGAGCTCAAGCTC 722

RESULT 4 B0864461

LOCUS B0864461 720 bp mRNA linear EST 14-AUG-2002
 DEFINITION QG26M12.yg.ab1 QG-ABCDI lettuce salinas Lactuca sativa cDNA clone
 QG26M12, mRNA sequence.

ACCESSION B0864461

VERSION B0864461.1 GI:22249926
 KEYWORDS EST.

SOURCE Lactuca sativa

ORGANISM Lactuca sativa
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
 Cichorieae; Lactuca.

REFERENCE AUTHORS

1 (bases 1 to 720)
 Kozik.A., Michelmore.R.W., Knapp.S., Matvienko.M., Rieseberg.L.,
 Lin.H., van Damme.M., Lavelle.D., Chevalier.P., Ziegler.J.,
 Ellison.P., Kolman.J., Slabaugh.M.S., Livingston.K., Zhou.Y.,
 Lai.Z., Church.S., Jackson.L. and Bradford.K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/
 Unpublished (2002)

TITLE JOURNAL

COMMENT Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmudson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 singleton, see http://cgdb.ucdavis.edu/ for details.

Plate: QG26 row: M column: 12.
 Location/Qualifiers

FEATURES source

1..720
 /organism="Lactuca sativa"
 /mol_type="mRNA"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="QG26M12"
 /lab_host="E.coli"
 /clone_lib="QG-ABCDI lettuce salinas"
 /note="Vector: pBRCNASTAB; the library was constructed
 from 10 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize


```

Db      424 AGTACTGAAGAGAGGCTGCTGAAGCATACACATAGCTGGATAAAGTTGAGAGTCTC 483
QY      1209 ACCGCACTGACTTAATCTCCACATGAACAGATACACAGTTAAAGCAATCTCGAAAGCCCT 1268
Db      484 AACGCTGTACAAACTTTGACATGAGCGCTACGACGTGAAGCCATCTTTGAAAGCAAC 543
QY      1269 AGTCTTCCCTAT---TGGTAGCGCCGCAAAAGCTCTCAAGGAGGCTAA 1312
Db      544 ACTCTCCCAATAGGAGGAGGAGCTGCAAAAGCGTCTGAAGAAGCTCA 590

RESULT 6
CF807326
LOCUS   CF807326
DEFINITION pSHB023xF07f USDA-IRAFS:Expression of Phytophthora sojae genes during infection and propagation Phytophthora sojae cDNA clone sHB023F07 5, mRNA sequence.
ACCESSION CF807326.1 GI:37995737
VERSION   CF807326.1
KEYWORDS  EST.
SOURCE    Phytophthora sojae
ORGANISM  Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.
REFERENCE 1 (bases 1 to 748)
AUTHORS   Tyler,B.
TITLES    Tyler,B. Not Published
JOURNAL   Unpublished (2003)
COMMENT    Contact: Tyler B
          Tyler lab
          VBI
          1880 Pratt Dr., Blacksburg, VA 24061, USA
          Tel: 540-231-7318
          Email: bmtyle@vt.edu
PCR PRIMERS
FORWARD: BK reverse primer
BACKWARD: BK reverse primer
Plate: 023 row: F column: 07
Seq primer: BK reverse primer
High quality sequence stop: 748.
          Location/Qualifiers
FEATURES             source
1..748
/organism="Phytophthora sojae"
/mol_type="mRNA"
/db_xref="taxon:67593"
/clone="SHB023F07"
/tissue_type="mycelium"
/cell_line="F6437"
/dev_stage="48 hr. post infection stage"
/lab_host="Soybean plant"
/clone_lib="USDA-IRAFS:Expression of Phytophthora sojae genes during infection and propagation"
/note="Vector: pBK-CMV, Site_1: EcoR1; Site_2: Xho1"

ORIGIN
Query Match 17.4%; Score 351.4; DB 14; Length 748;
Best Local Similarity 77.5%; Pred. No. 3.9e-60;
Matches 424; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY      700 AAGCTACACCGAGAAACATTGAGAGTTTGGACAGAGGAGCTCTATATACCGCGGTG 759
Db      202 AACCTCTCCAAAGAAACCGTGCACCTTCGCGCAACGCACCTCCATCTACCGCGGCG 261
QY      760 TTCAAGGCATCGGTGACAGAGAGATATAGCGCATTTATGGGATATAGTTGTAAAA 819
Db      262 TCACCCGACATAGATGACGCGGAAGATACGAAGCTCATCTATGGGACAATAGTTGTAGAA 321
QY      820 GAGAGGCCAAACCGCAAGAGGAGCAAGCTTTATTTGGGAGGTTATGACAAGAGAGAAA 879
Db      322 GAGAAGGCCAAACGAGGAAGGAGCAAGTTTACCTGGGTGTTTATGACAGAGACATA 381
QY      880 AAGCAGCTAGGCGTATGATTATAGCCGCACTCAAGTATTGGGGAACCAACCACTACTACTA 939

```

```

Db      382 AGCAGCCGAGGCTTACGATCTCGCAGCTCTCAAGTACTGGGTCCAACTACCACCACCA 441
QY      940 ACTTCCCATGAGCGCAATATGAAAGAGGTAGAGAGATGAAGCATGACAGGCAAG 999
Db      442 ACTTCCCATTTCCAACTATGAGAGGAACTGGAGAGATGAAGAAACATGACAGGCAAG 501
QY      1000 AGTATGTGCTCCTCAGTCGCGCAGGAAAAGTAGTGTCTCTCTCGTGGTGCATCGAATTATC 1059
Db      502 AGTTGTGTGCTTCTCTACGAAGGAGAGCAGTGTCTCTCTAGGGGGGCTCTATATACA 561
QY      1060 GTGAGGTAAACAGACATCACCAACATGGAAGATGGCAAGCTAGGATAGGAAGAGTGCCTG 1119
Db      562 GAGGAGTGAAGACACACCAACAGCATGGCGGAGTGGCAGGCGAGATAGGCAGAGTTGCCG 621
QY      1120 GTACAAAGAGACCTCTACTTTGGGAACCTTTGGCACACAAGAGAGCTGCAGAGGCATACG 1179
Db      622 GAAACAAAGAGACCTCTACTCTTGGAACTTTTACGACCCCAAGAGAGCTGTGAGGCCCTATG 681
QY      1180 ACATTGGCGGCATCAAAATTCAGAGGATTAACCCAGTACTAACTTCGACATGAACAGAT 1239
Db      682 ACATTGCTGTATCANATTCAGGGGATTAATGCACTCANCAACTTTGACATGATCGCT 741
QY      1240 ACAAGCT 1246
Db      742 ACGATGT 748

RESULT 7
BF647766
LOCUS   BF647766
DEFINITION BF647766 674 bp mRNA linear EST 20-DEC-2000 clone NF025G09EC 5', mRNA sequence.
ACCESSION BF647766
VERSION   BF647766.1 GI:11912896
KEYWORDS  EST.
SOURCE    Medicago truncatula (barrel medic)
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 674)
AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLES    Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research
JOURNAL   Unpublished (2000)
COMMENT    Contact: Dixon RA
          Plant Biology Division
          The Samuel Roberts Noble Foundation
          2510 Sam Noble Parkway, Ardmore, OK 73402, USA
          Tel: 580 221 7302
          Fax: 580 221 7380
          Email: radixon@noble.org
          Insert Length: 674 Std Error: 0.00
          Plate: 025 row: G column: 09
          Seq primer: TCACACAGGAACACAGCTATGAC.
FEATURES             source
1..674
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF025G09EC"
/tissue_type="Cell cultures derived from root tissues"
/dev_stage="Cell suspensions were subcultured every 14 days. Cells were induced six days after subculture"
/clone_lib="Elicited cell culture"
/note="Vector: Lambda Zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."

```


RESULT 9

LOCUS
DEFINITION
AW200688 644 bp mRNA linear EST 02-DEC-2001
se92c07.y1 Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl027-229 5' similar to TR:Q41832 Q41832 OPEN READING FRAME. ;
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AW200688
AW200688.1 GI:6481417
EST
Glycine max (soybean)
Glycine max
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 644)

REFERENCE
AUTHORS
Shoenmaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,D., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)

TITLE
JOURNAL
COMMENT
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estowatson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 968 Std Error: 0.00
High quality sequence stop: 418.
Location/Qualifiers
1..644
/organism="glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-229"
/tissue.type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH103"
/clone_lib="Gm-cl027"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First-strand synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer [GAGCAGAGAGAGAGACTAGTCGAG(r).is] to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in
with cloned pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The XhoI site within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500 bp cutoff,
using GibcoBRL Life Technologies' cDNA Size Fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript(tm) II XR predigested vector
(pbluescript II SK+) that has been digested with EcoRI
(pbluescript II SK+).

FEATURES
source

ORIGIN

Query Match 16.9%; Score 339.4; DB 10; Length 644;
Best Local Similarity 77.8%; Pred. NO. le-57;
Matches 409; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 700 AACGTACACCGAAGAAGAACTATTGAGAGTTTGGACAGAGGACGCTATATACCCGGTG 759
Db 118 AACCTTCTCAAGAAAACCGTCGACACCCTTGGCCAAGCACCTCCATCTACCGCGCG 177
QY 760 TTACAAGGCATCGGTGGACAGGAAGATATAGGCCACATTTATPGGATTAATAGTTGTAATA 819
Db 178 TCACCCGACATAGTAGGACGGGAAGATAGCAAGCTCATCTATGGGACAAATAGTTGTAGAA 237
QY 820 GAGAAAGCCCAAACGCGCAAGGAAGACAAGTTTATTTGGGAGGTTATGACAAAGAGAAA 879
Db 238 GAGAAGGCCCAAACGCGCAAGGAAGACAAGTTTATCTCTGGTGGTATGACAAAGGAATA 297
QY 880 AAGCAGCTAGGGGCTTATGATTAGCCGCACTCAAGTATTGGGGAACCACTACTACTA 939
Db 298 AGGCAGCCAGGCTTACGATCTCGAGCTCTCAAGTACTGGGTCCAACTACCAACCA 357
QY 940 ACTTCCCCTAGGGGAATATGAAAAAGAGGTAGAGAGATGAAGCACATGACAAGGCAAG 999
Db 358 ACTTCCCATTTCCAACATATGAGAAGAACTGGAGAGATGAAGAACATGACAGGCAAG 417
QY 1000 AGTATGTTGCTCACTGCGCAGCAAAAAGTAGTGGTTTCTCTCGTGGTGCATCGATTATC 1059
Db 418 AGTTTGTGCTTCTTACGAGGAGAGGACAGTGGTTTCTTAGSGGGGCTCTATATACA 477
QY 1060 GTGAGTAAACAAGACATCACCAACATGGAAGATGGCAAGCTAGGATAGGAAGAGTCGCG 1119
Db 478 GAGGAGTGAGGAGACACCAACAGCATGGCGGATGGCAGCGAGAATAGCAGAGTTGCGG 537
QY 1120 CTACAAGACCTCTACTCTGGGAACCTTTTGGCACACAAGAAGCTGCAGAGGCATAGC 1179
Db 538 GAACAAGACCTCTACTCTTGGAACTTTAGCAACCCCAAGAGAGCTGCTGANGCCTATG 597
QY 1180 ACATTGGCGGCATCAAAATTCAGAGGATTAACCCAGTACTAATT 1225
Db 598 ACATTGCTGCTATCAAAATTCAGGCGATTAATTCAGTACCAAACT 643

RESULT 10
CK267021 882 bp mRNA linear EST 12-DEC-2003
CK267021
LOCUS
DEFINITION
ES7713099 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POAC822 5' end, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CK267021.1 GI:39823999
EST
Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Ascomycota; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 882)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Buell,C.R., Hart,A., Zismann,V., Karanycheva,S.A. and Baker,B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: EST713100
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

FEATURES

```

source
1. .882
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="FOAC822"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d,
and 2d). Set 2 were grown under the standard conditions and
then water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN
Query Match 16.8%; Score 339.2; DB 14; Length 882;
Best Local Similarity 72.2%; Pred. No. 1e-57;
Matches 456; Conservative 0; Mismatches 173; Indels 3; Gaps 1;

QY 673 AAGCGAGACATTTGATGATAGCTTGAAGCTACACCGAAGAAACTATTGACAGTTTGG 732
DB |||||
QY 116 AAAAGCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175
DB |||||
QY 733 GACAGAGACGCTCTATATACCGCGTGTTCACAGGCTCGGTGCACAGGAAGATATGAGG 792
DB |||||
QY 176 GGCAAGAACTTCAATCTACAGAGGTGTACAGACATAGTACAGAGGAAGATATGAAG 235
QY 793 CACATTTATGGGATAATAGTTGTGTAAGAGAGAGCGCAACCGCAAGGAAGCAAGTTT 852
DB |||||
QY 236 CTCATCTATGGGATAATAGTTGTGTAAGAGAGAGCGCAACCGCAAGGAAGCAAGTTT 882
QY 853 ATTTGGGAGGTTATGACAAAGAGAAAGACAGCTAGCGGTATATGATTTAGCGGCACTCA 912
DB |||||
QY 296 ACTTGGTGGATATGACAAAGAGATATAGCAGCAAGGCGATATGACTTGGCAGCTCTTA 355
QY 913 AGTATTGGGAAACACCACTACTACTTAATCTCCCATAGCGAATATGAAAAAGAGGTAG 972
DB |||||
QY 356 AGTATTGGGTTCCACAGCTACCACTCCCTGTTACTAATATTACTAAAGAAATGG 415
QY 973 AAGAGATGAAGCATCATCAGAGGCAAGATATGTTCCTCCTCAGTCCGCGCAAGAGTAGTG 1032
DB |||||
QY 416 AGGAATGAACACATGACTTACGAGAAATCATTTGCTCTTAAGAGAAAGTAGTG 475
QY 1033 GTTTCTCTCGTGGTCATCGATTATCGTGGAGTAAACAGACATCAACATCGGAAGAT 1092
DB |||||
QY 476 GTTTCTCTCGGAGGCTTCGATTACCGGGTGTGACAGGATCATCAACAGGCGCT 535
QY 1093 GGAAGCTAGGATAGGAGAGTCCCGGTAAACAGACTCTACTTTGGGAACTTTTGGCA 1152
DB |||||
QY 536 GGCAGCGGAATTTGGCCGAGTGTCTGGAACAGAGATCTTTACCTTGGAACTTTGCTA 595
QY 1153 CACAGAGAGAGCTGCAGAGGCAATGACATTCGCGCCCATCAATTCAGAGGATTAACCG 1212
DB |||||
QY 596 CTGAGGAGGAGAGCGAGCGAGCTGATGACATTCAGCCATTAAGTTTCAGGGAGTGAATG 655

```

```

QY 1213 CAGTGACTTAACCTCGCATGACAGATACAGGTTAAAGCATCTCTCGAAGCCCTAGTC 1272
DB |||||
QY 656 CAGTGACCAACTTCGAGATGAATCGTTATGATGTGAGGCGCATCATGCAAG---TTCCC 712
DB |||||
QY 1273 TTCTATTGTTAGTACGCGCGCAAAACGTTCTCAAG 1304
DB 713 TCCAGTTGTTGTAACCGCTTAAGCGTTTGAAG 744

AW780688 500 bp mRNA linear EST 03-DEC-2001
sl75e07.y1 Gm-c1027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1027-7165 5' similar to TR:Q41832 Q41832 OPEN READING FRAME. ;
mRNA sequence.

AW780688
AW780688.1 GI:7795291
EST.
Glycine max (soybean)
Glycine max
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 500)
Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 412.
Location/Qualifiers
1..500
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1027-7165"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/clone_lib="Gm-c1027"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First- stranded synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer [GAGAGAGAGAGAGAGAGACTAGTCGAG(T)18] to anchor
the primer at the 5' end of the poly(A) tract. After
second- strand synthesis, the cDNA ends were filled in
with cloned Pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The XhoI site within the

```

first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript (tm) II XR predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=30). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

ORIGIN

Query Match 16.7%; Score 335.4; DB 10; Length 500;
Best Local Similarity 79.7%; Pred. No. 6.7e-57;
Matches 396; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

736 AGAGGACGCTCTATATACCGCGGTGTACAAAGCATCGTGGACAGAGATATAGGAC 795
Db AGCGCACCTCCATCTACCGCGGTGTACCGGACATAGATGGACGGAGATACGAAGCTC 63

796 ATTATGGGATATAGTTTAAAGAGAGGCGCAACGCGCAAGGACAGAGTTTATT 855
Db ATCTATGGGACATAGTTTGAAGAGAGGCGCAAGGAGGAGAGAGAGTTTACC 123

856 TGGGAGTTTATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 915
Db TGGTGGTTATGACAGAGAGATAGGACGAGCGGCTTACGATCTCGAGCTCTCAAGT 183

916 ATTGGGAGACCACTACTACTTATCTCCCATGAGCGAATATGAAAGAGGTAGAG 975
Db ACTGGGTCCTCACTACCAACCACTTTCCCACTTATGAGAGAGAGAGAGAGAG 243

976 AGATGAGAGACATGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1035
Db AGATGAGAGACATGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 303

1036 TCTCTCGTGTGCATCGATTTATCGTGGAGTAAAGAGATCAACAGATCAACAG 1095
Db TCTCTAGGGGGGCTCTATATACAGAGAGAGTACGAGACACCAACAGCATGGC 363

1096 AAGCTAGGATAGAGAGAGTGGCGGTAAAGAGAGAGAGAGAGAGAGAGAGAG 1155
Db AGCGGAGATAGAGAGAGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423

1156 AAGAGAGAGCTGAGAGAGATACGACATTCGGGCGCATCAATTCAGAGATTAA 1215
Db AAGAGAGAGCTGAGAGAGCTGAGAGAGATACGACATTCGGGCGCATCAATTC 483

1216 TGACTAATCTGCATG 1232
Db TCACAACTTTGACATG 500

RESULT 12

BM307997
LOCUS 593 bp mRNA linear EST 02-JAN-2002
DEFINITION sak40e09.y1 Gm-c1036 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1036-5586 5' similar to TR:Q9SYC2 Q9SYC2 F11M15.6 PROTEIN. 1;
mRNA sequence.

ACCESSION

BM307997

VERSION

BM307997.1 GI:18039703

KEYWORDS

EST.

SOURCE

Glycine max (soybean)

ORGANISM

Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE

1 (bases 1 to 593)
Shoemaker, R., Keim, P., Vodkin, J., Erpelding, J., Coryell, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Watterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

TITLE

JOURNAL

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 427.

Location/Qualifiers

1..593

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1036-5586"

/tissue_type="somatic embryos cultured on MSD 20"

/lab_host="DH10B"

/clone_lib="Gm-c1036"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This

cDNA library was constructed from mRNA isolated from

somatic embryos (age ranging from 2 months to 9 months)

cultured on MSD 20. The library was prepared using the

Life Technologies pSuperScript cDNA library construction

kit. Complementary DNA was synthesized from mRNA using a

poly (dT) sequence with a NotI restriction site. SalI

linkers adapters were ligated to the blunt-ended cDNA

fragments followed by NotI digestion. The cDNA fragments

were directionally cloned into the NotI-SalI restriction

site of the pSPORT1 vector. The ligated cDNA fragments

were transformed into E.coli Electromax DH10B host cells.

This library was constructed in the laboratory of Dr. Lila

Vodkin by Anu Khanna at the University of Illinois at

Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Query Match 16.6%; Score 335.2; DB 12; Length 593;
Best Local Similarity 77.5%; Pred. No. 7.1e-57;
Matches 406; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

681 ACTATTGATGATAGCGTTGAAGCTACACCGAAGAAACTATTGAGAGTTTGGACAGG 740

Db ACTACTACTACTACTGTTGAAGCTGCACCTAGAAGACATTGGATACATTCCGGCAGAG 125

741 ACGTCTATATACCGCGGTGTACAGGCATCGTGGACAGGAAGATATAGGCGACATT 800

Db ACATCCATATATCGTGGAGTAAACGCGACATAGTGGACTGGAAGGTATGAAGCTCACCT 185

801 TGGGATATATAGTGTAAAGAGAGAGCGCAAAACCGCGCAAGAAAGACAGTTTATTGGGA 860

Db TGGGATATATAGTGTAGAGAGAGAGCGCAATCAAGAAAGAGCGCAAGTTTATTGGGT 245

861 GGTATGACAAAG 920

Db GGATATGATAAAG 305

921 GGAACCCACTACTACTACTTCCCATGAGCGGATATGAAGAGAGAGAGAGAGATG 980

Db GGAACCATCCACACTTACCACCTTCCCAATTAGTACTATGAGAGAGAGAGAGATG 365

981 AAGCACATGACAAAGAGAGAGATGTTGCTCTCCTCAGTCCGACAGAAAGTAGTGTTCTCT 1040

Db AAACACATGACGCGCAAGAGAAATTTGCTGCTCCATTAGAGAGAAAGAGAGTGTCTCC 425

Db 575 CGACCGCTGCTCCGATGAATCTTTAGCCATCCAAACATGTTCCCTTCTCCCTTTGGTG 634

Qy 301 TCGTGTGATGCTTTTACACGAGACAAATAGTCACTCCCGAGATT 348

Db 635 TCGTCTGATGCTTTTACACGAGACAAATAGTCACTCCCGAGATT 682

RESULT 15
CD475882
LOCUS

DEFINITION
CD475882 545 bp mRNA linear EST 04-JUN-2003
mRNA sequence.
CD475882
CD475882.1 GI:31397150
EST.
Nuphar advena
Nuphar advena
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; stem Magnoliophyta; Nymphaeaceae;
Nuphar.
1 (bases 1 to 545)
dePamphilis,C., Carlson,J., Ma,H., Frohlich,M., Tanksley,S.,
Leebens-Mack,J., Field,D., Arrington,J., Zahn,D., Kong,H.,
Druckemiller,M., Landherr,L., Hu,Y., Ilut,D., Wall,K.,
Plock,S.Chloresan,S., Albert,V., Doyle,J., Miller,W.,
Oppenheimer,D., Soltis,D., Soltis,P. and Theissen,G.
Generation of ESTs from early flower buds of Nuphar advena
Unpublished (2002)
Contact: Claude dePamphilis or James Leebens-Mack
Mueller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 8413
Fax: 814 865 9131
Email: cwd3@psu.edu or jhl10@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (<http://pgn.cornell.edu>)
Plate: nad03-15msl row: h column: 07
Seq primer: M13F.

FEATURES
source
1..545
/organism="Nuphar advena"
/mol_type="mRNA"
/db_xref="taxon:77108"
/clone="nad03-15msl-h07"
/tissue_type="flower buds"
/dev_stage="c= 2.5mm"
/lab_host="SOLR"
/clone_lib="nad03"
/notes="Vector: pBluescript SK (-/-); Site 1: EcoRI;
Site 2: XhoI; Only floral buds with diameter of 2.5 mm of
less were used for RNA isolation. This is a directionally
cloned, non-normalized library. Avg. insert length: 1134;
Primers: M13F and M13R; Antibiotic: 50 ug/ml Ampicillin;
Primary titer: 2E6 pfu total; Amplified titer: 3.2E10
pfu/ml; Mass Excised Titer: SE10 total; This library has
been generated by the Floral Genome Project (FGP). The
Floral Genome Project is funded by NSF's Plant Genome
Research Program (DBI-0115684). More information about the
project can be obtained at <http://fgp.bio.psu.edu>"

ORIGIN
Query Match 16.1%; Score 323.6; DB 14; Length 545;
Best Local Similarity 75.1%; Pred. No. 1.5e-54;
Matches 404; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

Qy 703 CTACCCAGAAACATTTAGAGTTTTCGACAGAGGAGCTCTATATACCGCGGTGTTA 762

Db 5 CTATGCATAGGAGTCCATTGACACCTTTGGACAGAGAACATCTTCAGTATAGAGGGGTTA 64

Qy 763 CAAGGCATCGTGCAGGAGAGATATGAGGCACATTTTGGGATATATAGTTGTAAAGAG 822

Db 65 CAAGACACAGATGGACTGTAGATATGAAGCCCATCTTTGGGACATAGCTGTAGAGG 124

Qy 823 AAGGCCAAACCGCCAAAGGAAGCAAGTTTATTGGGAGGTTATGACAAAGAAAGAAAG 882

Db 125 AAGGGCAAAACCAAGAAAGGAAGCAAGTCTATCTGGGTGGCTATGACATGGGAAGAAAG 184

Qy 883 CAGCTAGGGCTTATGATTTAGCCGCACTCAAGTATTGGGGAACCCACCTACTACTAACT 942

Db 185 CAGCTAGGGCTTATGACTTTAGCAGCACTCAATATTTGGGGCTGTCTACTCACAAT 244

Qy 943 TCCCATGAGCGCAATATGAAAAAGAGGTAGAAGAGATGAAGCAATGACAAAGCAAGAGT 1002

Db 245 TTCCATTGAGCAATTACCAGGAAGAACTTGAAGGGATGAAGAACATGAGCAGCAAGAGT 304

Qy 1003 ATGTTGCTCAGTCCGCGCAGGAAAGTAGTGTCTCTCTGCTGCTGATCGATTATCGTG 1062

Db 305 ATGTTGCTCAGTTAAGGAGGAAAGTAGTGGGTTTCAAGAGCGGATCTATCTATAGAG 364

Qy 1063 GAGTAAACAAGACATCACCAACATGGGAGATGGCAAGCTAGGATAGGAAGAGTCCCGGTA 1122

Db 365 GTGTAACAAGGCATCATCAACATGGGAGGTGGCAAGCTAGGATTGGTAGAGTAGTGGAA 424

Qy 1123 ACAAGACCTCTACTTGGGAACTTTTGGCACACAAGAAAGAGCTGCAGAGGCATACGACA 1182

Db 425 ACAAGACCTTTTATCTTGGGACCTTTAGTACTCAGGAGGAGGAGCTGGAAGCATACGCG 484

Qy 1183 TTGGGGCCATCAAAATTCAGAGGATTAACCCGAGTGAATCTTCGACATGAACAGATA 1240

Db 485 TTGCTGCAATCAAAATTCGAGGGTTGATGCTGTGACCACTTTGACATAACAGATA 542

Search completed: July 19, 2004, 10:20:06
Job time : 5297 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 05:43:48 ; Search time 156 Seconds
(without alignments)
7164.558 Million cell updates/sec

Title: US-09-980-364-1

Perfect score: 2014

Sequence: 1 gttcatctcttcttcttaag.....gttgataaatttcagact 2014

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : 1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
6: /cgn2_6/prodata/2/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	330.6	16.4	2148	4	US-09-227-421-1
2	330.6	16.4	2148	4	US-09-479-855-1
3	324.8	16.1	1738	4	US-09-227-421-4
4	324.8	16.1	1738	4	US-09-479-855-4
5	129	6.4	1680	2	US-08-700-152A-3
6	79.8	4.0	301	4	US-09-313-294A-6727
7	67.2	3.3	7218	1	US-08-232-463-14
8	55.2	2.7	294	4	US-09-313-294A-6087
9	55	2.7	7218	1	US-08-232-463-14
10	44.6	2.2	5562	4	US-10-204-708-63
11	44.4	2.2	1024	4	US-09-328-475C-50
12	43.4	2.2	8607	4	US-10-204-708-72
13	43.2	2.1	658	3	US-08-998-416-595
14	43.2	2.1	6040	4	US-10-204-708-69
15	42.8	2.1	11721	4	US-09-026-039-3
16	42.8	2.1	11721	4	US-09-026-039-3
17	42.6	2.1	6368	4	US-10-204-708-68
18	42.4	2.1	396	4	US-09-640-173-53
19	42.4	2.1	396	4	US-09-713-550-53
20	42.4	2.1	731	1	US-08-451-405A-2
21	42.2	2.1	240	1	US-08-628-417-6
22	42.2	2.1	5535	4	US-10-204-708-18
23	42	2.1	5152	4	US-10-204-708-47
24	41.8	2.1	674	4	US-09-620-405B-465
25	41.8	2.1	674	4	US-09-433-826B-465
26	41.8	2.1	674	4	US-09-604-287A-465
27	41.8	2.1	674	4	US-09-834-759-465

28	41.8	2.1	8537	4	US-10-204-708-41	Sequence 41, Appl
29	41.4	2.1	1117	3	US-09-247-373B-33	Sequence 33, Appl
30	41.4	2.1	2447	2	US-09-014-969-14	Sequence 14, Appl
31	41.4	2.1	2634	4	US-09-463-238-3	Sequence 3, Appl
32	41.4	2.1	6659	4	US-10-204-708-6	Sequence 6, Appl
33	41.2	2.0	1378	4	US-09-149-476-208	Sequence 208, App
34	41.2	2.0	6583	4	US-10-204-708-26	Sequence 26, Appl
35	40.8	2.0	174	4	US-09-302-161B-21	Sequence 21, Appl
36	40.8	2.0	1048	4	US-09-489-847-38	Sequence 38, Appl
37	40.8	2.0	1405	4	US-09-202-161B-6	Sequence 6, Appl
38	40.6	2.0	2323	4	US-09-149-476-24	Sequence 24, Appl
39	40.6	2.0	7286	3	US-09-331-581-3	Sequence 3, Appl
40	40.6	2.0	7938	3	US-09-331-581-14	Sequence 14, Appl
41	40.6	2.0	8961	4	US-10-204-708-80	Sequence 80, Appl
42	40.6	2.0	19124	2	US-08-487-826B-13	Sequence 13, Appl
43	40.4	2.0	1558	1	US-08-467-607-2	Sequence 2, Appl
44	40.4	2.0	1558	2	US-08-469-362-2	Sequence 2, Appl
45	40.4	2.0	1558	2	US-08-850-392-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-227-421-1
; Sequence 1, Application US/09227421
; Patent No. 6559357
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
; FILE REFERENCE: 023070-090700PC
; CURRENT APPLICATION NUMBER: US/09/227,421
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 09/227,421
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2148
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (269)..(1936)
; OTHER INFORMATION: AINTEGUMENTA (ANT)
; NAME/KEY: misc_feature
; LOCATION: (1109)..(1339)
; OTHER INFORMATION: encodes first AP2 domain
; NAME/KEY: misc_feature
; LOCATION: (1340)..(1414)
; OTHER INFORMATION: encodes linker region
; NAME/KEY: misc_feature
; LOCATION: (1415)..(1621)
; OTHER INFORMATION: encodes second AP2 domain
US-09-227-421-1

Query Match	16.4%	Score 330.6;	DB 4;	Length 2148;
Best Local Similarity	75.4%	Pred. No. 9.7e-85;		
Matches 411;	Conservative	0;	Mismatches 134;	Indels 0;
Gaps 0;				
QY	714	AAACTATTGAGATTTTGGACAGAGCGTCTATATACCGCGGTGTACAGGCAATCGG	773	
Db	1082	AAATCTATCGTACTTTTGGACAGAGCTTCTCAATACCGAGCGGTTCACAGACATAGA	1141	
QY	774	TGGACAGGAAGATATGAGGCACATTTATGGATATAGTTCTATAAAGAGAGGCCAAACG	833	
Db	1142	TGGACTGTAGTATGAGCTCTATCTATGGACATAGTTTCACAGAGGAGGTACACAGT	1201	
QY	834	CCCAAAGAGAACAAAGTTTATTTGGAGGTATGCAAGAGAAAGAGAGCTTAGGCT	893	

Db 1202 AGAAAGGAGACACAGTTTATCTGGAGGTTATGATATGGAGGAGAAAGCTGCTCGAGCA 1261
Qy 894 TATGATTTAGCCGCACTCAAGTATTGGGAAACCACCACTACTACTAACTTCCCATGAGC 953
Db 1262 TATGATCTTGCTGCACTCAAGTACTGGGTCCTCTACTACACCAATTTCTCTGGGAG 1321
Qy 954 GAATATGAAGAAGGTAGAGATGAGACATGACATGACAGGCAAGAGTATGTTGCTCA 1013
Db 1322 AATATCAGAAGAGATTGAAGACATGAAGAACATGACTACAGAAGATATGTTGCACAT 1381
Qy 1014 CTGGCAGGAAAGTAGTGGTTTCTCTGCTGCTGATCGATTTATCTGCTGAGTAAACAAGA 1073
Db 1382 TTGAGAGGAAAGACAGTGGTTTCTCTAGGGTCTTCCATCTATAGAGAGTACACAAGA 1441
Qy 1074 CATCACCAATGAAAGATGCAAGATGAGTAGAGAGTGGCGGTACAAAGACCTC 1133
Db 1442 CATCACCAATGAAAGATGCAAGATGAGTAGAGTGGCGGTACAAAGACCTC 1501
Qy 1134 TACTTGGGAACCTTTTGGCACAAAGAAAGCTGCAGAGGCATACGACATTTGGGCGCATC 1193
Db 1502 TACCTTGGAACTTTTGGAAACCCAGAGAGCTGCAGAGCTTACGATGTAGCAGCAAT 1561
Qy 1194 AATTCAGAGGATTAACCGAGTGAATCTTGCACATGAACATACAAAGTTAAAGCA 1253
Db 1562 AAGTTCGCTGGCACAAATGCTGTGACTAACTTTGATATCAGAGGTACGATGTTGTCGT 1621
Qy 1254 ATCCCT 1258
Db 1622 ATCAT 1626

RESULT 2

US-09-479-855-1

; Sequence 1, Application US/09479855

; Patent No. 6639128

; GENERAL INFORMATION:

; APPLICANT: Fischer, Robert L.

; APPLICANT: Mizukami, Yukiko

; APPLICANT: The Regents of the University of California

; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility

; FILE REFERENCE: 023070-090720US

; CURRENT APPLICATION NUMBER: US/09/479,855

; PRIOR FILING DATE: 2000-01-07

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2148

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: AINTEGUMENTA (ANT) cDNA

; NAME/KEY: CDS

; LOCATION: (269)..(1936)

; OTHER INFORMATION: AINTEGUMENTA (ANT)

US-09-479-855-1

Query Match

Best Local Similarity 16.4%; Score 330.6; DB 4; Length 2148;

Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

Qy 714 AAAAATATTGAGAGTTTGGACAGAGAGCTCTATATACCGCGTGTACAGGCGATCGG 773
Db 1082 AATCTATCGATCTTTTGGACACAGAACTTCTCATACCGAGCGTTACAGACATAGA 1141
Qy 774 TGGACAGAGATATAGGACATTTATGGGATTAATAGTTGTAAGAGAGAGCCAAACG 833
Db 1142 TGGACTGGTATATGAAGTCTATCTATGGGCAATAGTTTCAAGAGGAGTCAAGT 1201
Qy 834 CGAAAGCAAGCAAGTTTATTTGGGAGGTTATGACAAAGAGAAAGAGCAGTGGGCT 893
Db 1202 AGAAAGGAGACAGTTTATCTGGAGGTTATGATATGAGGAGAGAAAGCTGCTCGAGCA 1261

Qy 894 TATGATTTAGCCGCACTCAAGTATTGGGAAACCACCACTACTACTAACTTCCCATGAGC 953
Db 1262 TATGATCTTGCTGCACTCAAGTACTGGGTCCTCTACTACACCAATTTCTCTGGGAG 1321
Qy 954 GAATATGAAGAAGGTAGAGATGAGACATGAGACATGACATGACAGAGATATGTTGCTCA 1013
Db 1322 AATATCAGAAGAGATTGAAGACATGAAGAACATGACTACAGAAGATATGTTGCACAT 1381
Qy 1014 CTGGCAGGAAAGTAGTGGTTTCTCTGCTGCTGATCGATTTATCTGCTGAGTAAACAAGA 1073
Db 1382 TTGAGAGGAAAGACAGTGGTTTCTCTAGGGTCTTCCATCTATAGAGAGTACACAAGA 1441
Qy 1074 CATCACCAATGAAAGATGCAAGATGAGTAGAGAGTGGCGGTACAAAGACCTC 1133
Db 1442 CATCACCAATGAAAGATGCAAGATGAGTAGAGTGGCGGTACAAAGACCTC 1501
Qy 1134 TACTTGGGAACCTTTTGGCACAAAGAAAGCTGCAGAGGCATACGACATTTGGGCGCATC 1193
Db 1502 TACCTTGGAACTTTTGGAAACCCAGAGAGCTGCAGAGCTTACGATGTAGCAGCAAT 1561
Qy 1194 AATTCAGAGGATTAACCGAGTGAATCTTGCACATGAACATACAAAGTTAAAGCA 1253
Db 1562 AAGTTCGCTGGCACAAATGCTGTGACTAACTTTGATATCAGAGGTACGATGTTGTCGT 1621
Qy 1254 ATCCCT 1258
Db 1622 ATCAT 1626

RESULT 3

US-09-227-421-4

; Sequence 4, Application US/09227421

; Patent No. 6859357

; GENERAL INFORMATION:

; APPLICANT: Fischer, Robert L.

; APPLICANT: Mizukami, Yukiko

; APPLICANT: The Regents of the University of California

; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility

; FILE REFERENCE: 023070-090700PC

; CURRENT APPLICATION NUMBER: US/09/227,421

; PRIOR FILING DATE: 1999-01-08

; PRIOR APPLICATION NUMBER: US 09/227,421

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1738

; TYPE: DNA

; ORGANISM: Brassica napus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1647)

; OTHER INFORMATION: BANT, ANT ortholog from Brassica napus (Canola)

US-09-227-421-4

Query Match

Best Local Similarity 16.1%; Score 324.8; DB 4; Length 1738;

Matches 413; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 714 AAAAATATTGAGAGTTTGGACAGAGAGCTCTATATACCGCGTGTACAGGCGATCGG 773
Db 790 AATCTATTGATCTTTTGGACACAGAACTTGGCAATACCGAGCGTTACAGACATAGA 849
Qy 774 TGGACAGAGATATAGGACATTTATGGGATTAATAGTTGTAAGAGAGAGCCAAACG 833
Db 850 TGGACTGGTATGATGAAGTCTATCTATGGGCAATAGTTTCAAGAGGAGTCAAGT 909
Qy 834 CGAAAGCAAGCAAGTTTATTTGGGAGGTTATGACAAAGAGAAAGAGCAGTGGGCT 893
Db 910 AGAAAGGAGACAGTTTATCTGGGAGGTTATGATATGAGGAGAGAGCTGCTCGAGCA 969
Qy 894 TATGATTTAGCCGCACTCAAGTATTGGGAAACCACCACTACTACTAACTTCCCATGAGC 953


```

Query Match          2.2%; Score 44.4; DB 4; Length 1024;
Best Local Similarity 45.2%; Pred. No. 0.012;
Matches 90; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1810 AAGGTGGTGGAAGTGGTCCCAACATTTTACAGTTTGGGACGACAATTAGAAAAATAGT 1869
DB 353 AAAANNNGNANVAGAGCNCNCCNNAANANNCNAAAAANNNTATANGATNNCCCT 294

QY 1870 TAAAGATCTTTTGTATATATCGGTGTGTGTGCTGTGTGACAGTGTGATCTTTGATTAT 1929
DB 293 AAAAGATTTTNNANNNGNNANCTTTACCTTAAAAATTTTTTANTNTNTTTTTTTTTTT 234

QY 1930 GTTTTTTCTTCTCTCTTTTCTTTTCTTTCTTGGTAAATTTCTTAAGACTTATTTTAGTTTC 1989
DB 233 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 174

QY 1990 CATTAGTGTGATAAAATTTT 2008
DB 173 TTTTTTTTTTTTTTTTTTTT 155

RESULT 12
US-10-204-708-72
; Sequence 72, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BRELIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 72
; LENGTH: 8607
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-72

Query Match          2.2%; Score 43.4; DB 4; Length 8607;
Best Local Similarity 57.9%; Pred. No. 0.077;
Matches 77; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1876 TCTTTAGTTTATGCGTTCGTCTGTCGGTGAACAGTGTACTTTGATATGTTTTT 1935
DB 6988 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7047

QY 1936 TCTTTCTCTTTTCTTTTCTTGGTTAAATTTCTTAAGACTTATTTTAGTTTCCATTAG 1995
DB 7048 TTTTTTTCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTATTTTGT 7107

QY 1996 TTGGATAAATTTT 2008
DB 7108 TTATAATTTTTT 7120

RESULT 13
US-08-998-416-595

```

Sequence 595, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYBII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cortwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998.416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 595:
SEQUENCE CHARACTERISTICS:
LENGTH: 658 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: FAG1408RP
IS-08-998-416-595

[illegible]

RESULT 14
US-10-204-708-69
; Sequence 69, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander

; APPLICANT: PIERPENEROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: By Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 69
 ; LENGTH: 6040
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-59

 Query Match 2.1%; Score 43.2; DB 4; Length 6040;
 Best Local Similarity 56.2%; Pred. No. 0.072;
 Matches 81; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

 QY 1865 AVAGTTAAGATCTTTAGTATATACGTTGTGTGTGGTGAACAGTGTGATCTTTC 1924
 Db 2015 AAATTATATCTTTTATATATATGTTGTTTGTATATATAGTATATAGATGG 2074

 QY 1925 ATATGCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1984
 Db 2075 AATTATTTATTTAAATTTTATTTTATTTGATTTGCAATTTTTCGAAGATATGAGTAT 2134

 QY 1985 GTTCCATTAGTTGGATAAATTT 2008
 Db 2135 GTTTGTATTTTAGATTTTTTTT 2158

 RESULT 15
 US-09-026-039-3
 ; Sequence 3, Application US/09026039
 ; Patent No. 6329567
 ; GENERAL INFORMATION:
 ; APPLICANT: Jofuku, K. Diane
 ; APPLICANT: Okamuro, Jack K.
 ; TITLE OF INVENTION: Methods for Improving Seeds
 ; NUMBER OF SEQUENCES: 103
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM: disk
 ; MEDIUM TYPE: Floppy
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/026,039
 ; FILING DATE: 19-FEB-1998
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/912,272
 ; FILING DATE: 15-AUG-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/879,827

```
/ FILING DATE: 20-JUN-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/700,152
/ FILING DATE: 20-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 023070-067230US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 11721 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: -
/ LOCATION: 1..11721
/ OTHER INFORMATION: /note= "Arabidopsis APETALA2 (AP2)"
/ OTHER INFORMATION: genomic sequence"
US-09-026-039-3
```

```
Query Match
Best Local Similarity 2.1%; Score 42.8; DB 4; Length 11721;
Matches 111; Conservative 0; Mismatches 92; Indels 3; Gaps 1;

QY 955 AATATGAAAAAGAGGTAGAGAGATGAAGCACATGCACAAGGCAAGAGTATGTTGCCTCAC 1014
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2035 AATATTGAAATTGAATTTTATAGATGACTAATTTAACCAAGGAGAGTTCGTACCGTAC 2094
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1015 TGGCAGCAAAAGTACGTGTTTCTCTCGTGTGTCATCGATTATTCGTGGAGTAAAGAGAC 1074
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2095 TTCGCCGACAAAGCACAGGCTCCCTCGAGGAGTTCGAAAGTATAGAGGTGTCACTTTGC 2154
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1075 ATCACCACATGGAAGATGGCAAGCTAGGATAGGAAGATCGCGCGTAAACAAAGACCTCT 1134
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2155 ATAAAGTG--TGGTCGTTGGGAAGCTCGAATGGGTCAATTCCTAGGCAAAAGTATAATT 2211
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1135 ACTTGGGAACCTTTGGCACACAGAA 1160
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2212 TCTCTCATTTTATATTCACCTCGAAA 2237
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 19, 2004, 10:22:49
Job time : 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 08:51:52 ; Search time 921 Seconds
(without alignments)

10668.347 Million cell updates/sec

Title: US-09-980-364-1

Perfect score: 2014

Sequence: 1 gttcattctcttcttaag.....gttgataaatttcagact 2014

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3190992 seqs, 2439311697 residues

Total number of hits satisfying chosen parameters: 6381984

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	420.6	20.9	2463	13	US-10-183-687-355 Sequence 355, App
2	413.6	20.5	1824	13	US-10-183-687-353 Sequence 353, App
3	413.6	20.5	2588	13	US-10-183-687-351 Sequence 351, App
4	385.8	19.2	1990	13	US-10-183-687-361 Sequence 361, App
5	381	18.9	2088	16	US-10-374-780A-1738 Sequence 1738, App
6	380.8	18.9	2169	13	US-10-424-599-91337 Sequence 91337, A
7	377.4	18.7	1818	13	US-10-225-066A-365 Sequence 365, App
8	377.4	18.7	1818	13	US-10-225-066A-855 Sequence 855, App
9	377.4	18.7	1818	13	US-10-374-780A-911 Sequence 911, App
10	377.4	18.7	1818	16	US-10-374-780A-333 Sequence 333, App
11	364.4	18.1	794	13	US-10-425-114-20762 Sequence 20762, A
12	364.4	18.1	806	16	US-10-374-780A-1742 Sequence 1742, App
13	357.4	17.7	1935	13	US-10-183-687-486 Sequence 486, App
14	357	17.7	1479	13	US-10-425-114-14859 Sequence 14859, A

15 357 17.7 2147 13 US-10-424-599-70525 Sequence 70525, A
16 352 17.5 2168 13 US-10-183-687-357 Sequence 357, App
17 347.8 17.3 1290 13 US-10-183-687-359 Sequence 359, App
18 347.8 17.3 1954 13 US-10-183-687-490 Sequence 490, App
19 337.8 16.8 1949 13 US-10-424-599-123485 Sequence 123485
20 335.2 16.6 1878 13 US-10-183-687-482 Sequence 482, App
21 334 16.6 1182 13 US-10-425-114-19324 Sequence 19324, A
22 334 16.6 1874 13 US-10-183-687-347 Sequence 347, App
23 334 16.6 2128 16 US-10-374-780A-1763 Sequence 1763, App
24 332.4 16.5 2116 13 US-10-424-599-100930 Sequence 100930, A
25 332.2 16.5 1231 13 US-10-425-114-11685 Sequence 11685, A
26 332.2 16.5 2052 13 US-10-225-066A-1013 Sequence 1013, App
27 332.2 16.5 2052 16 US-10-374-780A-2495 Sequence 2495, App
28 330.6 16.4 1422 15 US-10-059-911-26 Sequence 26, App
29 330.6 16.4 1668 15 US-10-059-911-38 Sequence 38, App
30 330.6 16.4 2148 15 US-10-059-911-34 Sequence 34, App
31 330.6 16.4 618 15 US-10-059-911-32 Sequence 32, App
32 330.2 16.4 1113 15 US-10-059-911-30 Sequence 30, App
33 330.2 16.4 1359 15 US-10-183-687-484 Sequence 484, App
34 330.2 16.4 1959 13 US-10-425-114-783 Sequence 783, App
35 330.2 16.4 2374 13 US-10-183-687-488 Sequence 488, App
36 327.6 16.3 1985 13 US-10-059-911-39 Sequence 39, App
37 327.6 16.3 1668 15 US-10-059-911-33 Sequence 33, App
38 327.2 16.2 2323 14 US-10-024-632-3 Sequence 3, App
39 327.2 16.2 1500 13 US-10-325-066A-591 Sequence 591, App
40 326.4 16.2 1500 16 US-10-374-780A-339 Sequence 339, App
41 326.4 16.2 1738 15 US-10-059-911-4 Sequence 4, App
42 324.8 16.1 1941 13 US-10-225-066A-381 Sequence 381, App
43 324.4 16.1 1941 16 US-10-374-780A-345 Sequence 345, App
44 324.4 16.1 1941 16 US-10-183-687-337 Sequence 337, App
45 319.4 15.9 1758 13 US-10-183-687-337 Sequence 337, App

ALIGNMENTS

RESULT 1

US-10-183-687-355
; Sequence 355, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shet, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; PRIOR FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 355
; LENGTH: 2463
; TYPE: DNA
; ORGANISM: Glycine max
US-10-183-687-355

Query Match 20.9%; Score 420.6; DB 13; Length 2463;
Best Local Similarity 76.4%; Pred. No. 1.8e-105;
Matches 516; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczynski, Mitchell C.
 ; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
 ; FILE REFERENCE: B51458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; PRIOR FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 351
 ; LENGTH: 2598
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-183-687-351

Query Match 20.5%; Score 413.6; DB 13; Length 2598;
 Best Local Similarity 79.7%; Pred. No. 1.6e-103; Indels 0; Gaps 0;
 Matches 488; Conservative 0; Mismatches 124;

QY	695	CGTTGAAGCTACCGGAGAACTATTGAGCTTTTGGACAGAGACCTCTATACCG	754
Db	773	CATTGAAGCTACCGGAGAACTATTGAGCTTTTGGACAGAGACCTCTATACCG	832
QY	755	CGGTGTTACAGGCTCGGTGGACAGGAGATATGAGGCACATTTATGGATAATAGTG	814
Db	833	TGGTGTAAAGGCTATAGTGGAGCGGAGGTACGAGCTCCCTGTGGATAATAGTG	892
QY	815	TAAAGAGAGGCGAAACCGCGAAAGAGCAAGTTTATTTGGGAGGTATGACAAAGA	874
Db	893	TAGAAGAGAGGCGAGACTCGCAAGAGGAGCAAGTTTACTTGGGTGTTATGATAAGA	952
QY	875	AGAAAAGAGCTAGGCTTATGATTTAGCGCACTCAAGTATTTGGGAGGAAACCACTAC	934
Db	953	AGAAAAGAGCTAGGCTTATGATTTAGCGCACTCAAGTATTTGGGAGGAAACCACTAC	1012
QY	935	TACTAACTTCCCATGAGCGAATATGAAAAGAGGTAGAGAGATGAAGCAGATGACAG	994
Db	1013	AACAAATTTTCCAATTAGCCACTATGAGAAAGAGTTGGAAGAAATGAAGCAGATGACAG	1072
QY	995	GCAGAGTATGTCCTCTGCTGCGAGGAAAGTATGTTCTCTCGTGGTGCATCGAT	1054
Db	1073	GCAAGAGTATGTCCTCTGCTGCGAGGAAAGTATGTTCTCTCGTGGTGCATCGAT	1132
QY	1055	TTATCGTGGAGTAACAAGACATCACCAACATGGAAGATGCGAAGCTTAGGATAGGAAGCT	1114
Db	1133	TTATCGGAGGTACGAGACACACCAACATGGAAGGTGCGAAGCTTAGGATAGGAAGCT	1192
QY	1115	CGCGGTAAACAGACCTCTACTTGGAACTTTTGGACACACAGAGAGAGCTGCGAGGC	1174
Db	1193	TGCTGGCAACAGATCTTACTTGGAACTTTTGGACACACAGAGAGAGCTGCGAGGC	1252
QY	1175	ATACGACATTTGGGCGCCATCAAAATTCAGAGATTAACCGCAGTGACTTAACCTTCGACATGAA	1234
Db	1253	ATATGATGAGAGCAATCAAAATTCGAGGACTAAGTGCTGTTTACAACTTTGACATGAG	1312
QY	1235	CAGTACAGCTTAAAGCAATCTCGAAGCCCTAGTCTTCTTATTTGTTAGCGCGCAAA	1294
Db	1313	CAGTATGAGCTGAAAGCATACTTTGAGAGCACCCTTTGCCAATAGTGGTGTCTGCAAA	1372
QY	1295	ACGTCTCAAGGA	1306
Db	1373	CGGTTTGAAGGA	1384

RESULT 4

US-10-183-687-361
 ; Sequence 361, Application US/10183687
 ; Publication No. US20030204870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve

; APPLICANT: Allen, William B.
 ; APPLICANT: Cahoon, Rebecca
 ; APPLICANT: Ebelbaum, Sabine
 ; APPLICANT: Farnodu, Omolayo O.
 ; APPLICANT: Harvell, Leslie T.
 ; APPLICANT: Jones, Todd
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Klein, Ted
 ; APPLICANT: Li, Changjiang
 ; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczynski, Mitchell C.
 ; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
 ; FILE REFERENCE: B51458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; PRIOR FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 361
 ; LENGTH: 1990
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-183-687-361

Query Match 19.2%; Score 385.8; DB 13; Length 1990;
 Best Local Similarity 76.9%; Pred. No. 7.2e-96; Indels 3; Gaps 1;
 Matches 484; Conservative 0; Mismatches 142;

QY	692	TAGCGTTGAAGCTACACCGAAGAACTATTGAGAGCTTTTGGACAGAGACCTCTATATA	751
Db	560	TACTGTTGAAGTTGCACTTAGAAGAACTTTGGATACATTCGGGCGACAGAACATCCATA	619
QY	752	CGCGGTGTTTACAGGCTCGGTGGACAGAGATATAGGCACATTTATGGATAATAG	811
Db	620	TCGTGGAGTAACTCGACATAGATGGACTGGAGGTATGAAGCTCATCTTTGGGATAATAG	679
QY	812	TTGTAAAGAGAGGCGCAAGCGCAAGAGAAAGCAAGTTTATTTGGGAGGTATGACAA	871
Db	680	CTGTAGAAGGAGGCGCAAGCGCAAGAGAAAGCAAGTTTATTTGGGATAATGATAA	739
QY	872	AGAAAGAAAGAGCTAGGCTTATGATTTAGCGGCACTCAAGTATTTGGGAGGATGACAA	931
Db	740	AGAAAGAAAGAGCTAGGCTTATGATTTAGCTGCACTGAAGTACTTGGGGGACATCCAC	799
QY	932	TACTACTTAACTTCCCATGAGCGAATATGAAAAGAGGTAGAGAGATGAAGCAGATGAC	991
Db	800	CACTACCAACTTTCCCAATTAGCAACTATGAGAAGGAATTTGGATGAATGAACACATGAC	859
QY	992	AAGCAAGATATTTGGCTCACTGCGCAGGAAAGTAGTGGTTTCTCTCGTGGTGCATC	1051
Db	860	GAGACAGAAATTTGTTGCGCCCAATTAGAAGGAAAGCAAGTGGTTTCTCCAGGGGTGCATC	919
QY	1052	GATTTATCGTGGAGTAACAAGACATCAACAATGGAAGATGGAAGCTTAGGATAGGAAG	1111
Db	920	AATGATATCGTGGAGTTACAGGCACTACCAACACGGAAGATGGAAGCAAGGATTGGCAG	979
QY	1112	AGTCGCGGTAAACAAGACCTCTACTTGGAACTTTTGGACACACAGAGAGAGCTCCAGA	1171
Db	980	AGTTGCAAGAAACAAAGATCTTTTACTTGGAACTTTTCTACTTGGAGAGAGAGCTGCAGA	1039
QY	1172	GGCATACGACATTTGGGCGCCATCAAAATTCAGAGATTAACCGCAGTGACTTAACCTTCACAT	1231
Db	1040	AGCATACGACATAGCAGCGATTAAGTTTCAAGGTCTCAACGCTGTGCACAACTTTGACAT	1099
QY	1232	GAACAGATCAAGCTTAAAGCAATCTCGAAGCCCTAGTCTTCTCTAT---TGGTAGCCG	1288
Db	1100	GAGCCGCTACGAGTGAAGCCCAATTTTGAAGCAACACTCTCCCAATAGGAGGAGCCG	1159
QY	1289	CGCAAAACGCTCTCAAGGAGGCTTAACCGTC	1317

Db 1160 TCGAAGCGCTCTGAAGAAGCTCAAGCTC 1188

RESULT 5

US-10-374-780A-1738

; Sequence 1738, Application US/10374780A

; Publication No. US2004001927A1

; GENERAL INFORMATION:

; APPLICANT: Sherman, Bradley K

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Heard, Jacqueline E

; APPLICANT: Haake, Volker

; APPLICANT: Cressman, Robert A

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Adam, Luc J

; APPLICANT: Reuber, T. Lynne

; APPLICANT: Keddle, James

; APPLICANT: Broun, Pierre E

; APPLICANT: Pilgrim, Marsha L

; APPLICANT: Dubell III, Arnold T

; APPLICANT: Pineda, Omaira

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

; FILE REFERENCE: MBI-0047 CIP

; CURRENT APPLICATION NUMBER: US/10374,780A

; CURRENT FILING DATE: 2003-02-25

; PRIOR APPLICATION NUMBER: 09/837,944

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 09/934,455

; PRIOR FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 10/225,066

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,067

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,068

; PRIOR FILING DATE: 2002-08-09

; NUMBER OF SEQ ID NOS: 2906

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1738

; LENGTH: 2088

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G1793

US-10-374-780A-1738

Query Match 18.9%; Score 381; DB 16; Length 2088;

Best Local Similarity 76.5%; Pred. No. 1.6e-94;

Matches 48; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 692 TAGCGTTGAAGCTACACCGAAGAACTATTGAGAGTTTGGACAGAGGACGCTCTATATA 751

Db 567 TACTGTTGAAGTTGCACCTAGAAGAACTTTGGATACATTCGGCGCAGAGAATCCATATA 626

QY 752 CCGCGGTGTACAGGCATCGTGGACAGGAATATGAGGCATTTATGGATATAG 811

Db 627 TCGTGGAGTAACCTGCATAGATGGACTGGAAGGTATGAAGCTCATCTTTGGGATAATAG 686

QY 812 TTGTAAGAGAGAGGCGCAACCGCGCAAGGAAGACAAAGTTTATTTGGAGGTTATGACAA 871

Db 687 CTGTAGAAGGAGGAGGCAATCAAGAAAGGAGCGCCAAAGTTTATTTGGGTGGATATATA 746

QY 872 AGAAGAAAGAGCTAGGCTTATGATTTAGCCGACCTCAAGTATTGGGGACCAACCAC 931

Db 747 AGAAGAGAAAGCAGCTAGGCTTATGATTTAGCTGAAAGTACTGGGGACATCCAC 806

QY 932 TACTACTAACTTCCCATCAGCGAATATGAAAAGAGGTAGAGAGATGAAGCAGATGAC 991

Db 807 CACTACCAACTTCCCATGATTAATCTATGAGAAGAAATGGATGAATGAACACATGAC 866

QY 992 AAGGCAAGAGTATGTTGCTCTACTCGCAGGAAAAAGTAGTGTCTCTCGTGTGTCATC 1051

Db 867 GCGACAAGAAATTTGTTGCTGCCATTAGAAAGAAAAGCAGTGTCTTCTCCAGGGTGCATC 926

QY 1052 GATTTATCGTGGAGTAAACAGACATCAACACATCGAAGATGGCAAGCTAGTAGGAAG 1111

Db 927 AATGTATCGTGGAGTTACAAAGGATCAACACACGGAAGATGGCAAGAAATGGCAG 986

QY 1112 AGTCGCGGTAAACAAAGACCTCTACTTGGGAACCTTTTGGCACACAAGAAAGCTGCAG 1171

Db 987 AGTTGCAGGAAACAAAGATCTTTACTTGGGAACCTTTCAGTACTGAAGAAGAGGCTGCTGA 1046

QY 1172 GGCATACGACATTTGGGCGCATCAAAATTCAGAGGATTAACCGAGTGAATACTTCGACAT 1231

Db 1047 AGCATACGACATAGCTGCGATAAAGTTTCAGAGGTCTCAACGCTGTCAAAAACCTTTGACAT 1106

QY 1232 GAACAGATACAACTTAAAGCAATCTCTGAAAGCCCTAGTCTTCTAT---TGGTAGGCG 1288

Db 1107 GAGCGCTAGCAGCTGAAGGCATCTTTGAAGGCAACACTCTCCCATAGGAGGAGGAGC 1166

QY 1289 CGCAAAAGCTCTCAAGGAGGCTAACCGTC 1317

Db 1167 TCGAAGCGCTCTGAAGAAGCTCAAGCTC 1195

RESULT 6

US-10-424-599-91337

; Sequence 91337, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 91337

; LENGTH: 2169

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_53488C.1

US-10-424-599-91337

Query Match 18.9%; Score 380.8; DB 13; Length 2169;

Best Local Similarity 75.8%; Pred. No. 1.9e-94;

Matches 485; Conservative 0; Mismatches 152; Indels 3; Gaps 1;

QY 681 ACTATTGATGATACGCTTGAAGCTACACCGAAGAACTATTGAGAGTTTGGACAGAGG 740

Db 568 ACTACTACTACTCTGTTGAAGTTGACCTTAGAAGACATTTGGATACATTCGGGCGAGAG 627

QY 741 ACCTCTATATACCGCGGTGTTACAGGCATCGGTGGACAGAAAGATATAGGCACATTTA 800

Db 628 ACATCCATATATCTGCGAGTAAACCGCACATAGATGGACTGGAAGGTATGAAGCTCATCT 687

QY 801 TGGGATATAGTTGTAAGAGAGAGGCGCAACCGCGCAAGGAAGACAAAGTTTATTTGGGA 860

Db 688 TGGGATATAGTGTAGAGAGGAGGCAATCAAGAAAGGAGCGCCAAAGTTTATTTGGGT 747

QY 861 GGTATACAAAGAGAAAAAGCAGCTAGGCTTATGATTTAGCCGACCTCAAGTATTGG 920

Db 748 GGATATGATAAGAGAAAAAGCAGCTAGGCTTATGATTTAGCTGCACTGAAGTACTGG 807

QY 921 GGAACCCACTACTACTACTTCCCATGAGCAATATCAAAAAGAGGTAGAGATG 980
 Db 908 GGGACATCCACCATCACTTCCAAATAGTAATCTATGAGAGAAATGGATGAATG 867
 QY 991 AAGCACATGACAAAGGCAAGATATGTTGCTCACTCGCGAGGAAAAGTATGTTTCTCT 1040
 Db 868 AAACACATGACGCGACAAAGAAATTTGTTGCTGCCATTAGAAAGGAAAAGCAGTGGTTTCTCC 927
 QY 1041 CGTGTGCATCGATTTATCGTGGAGTAAACAGACATCAACCAATGGAAGTGGCAAGCT 1100
 Db 928 AGGGTGATCAATGTTATCGTGGAGTTACAGGATCAACCAACAGGAGATGGCAAGCA 987
 QY 1101 AGATAGAAAGATGCGCGGTAAACAAAGACCTTACTTTGGAACTTTTGGCACACAGAA 1160
 Db 988 AGAATTGCGAGATGTCAGGAAACAAAGATCTTTACTTTGGAACTTTTCAGTACTGAAGAA 1047
 QY 1161 GAAGTCGAGAGGATAGCATGCGCCATCAAAATTCAGAGGATTAACCGCAGTACT 1220
 Db 1048 GAGGTGCTGAAGCATACGATAGCTGCGATTAAGTTCAAGGTCTCAACGCTGTCA 1107
 QY 1221 AACTTCGACATGAACAGATCAACAGTTTAAAGCAATCTCGAAAGCCCTAGTCTTCCAT- 1279
 Db 1108 AACTTTGACATGAGCGCTAGCAGGTGAAGCCATCTTTGAAAGCAACACTCTCCCAATA 1167
 QY 1280 --TGTAGCGCGCAAAACGCTCAAGAGGCTTAACCGTC 1317
 Db 1168 GGAGAGAGGTGCAAAAGCGTCTGAAAGAGCTCAAGCTC 1207

RESULT 7

US-10-225-066A-365

; Sequence 365, Application US/10225066A

; Publication No. US20030226173A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: RATCLIFFE, Oliver

; APPLICANT: RIECHMANN, Jose Luis

; APPLICANT: ADAM, Luc J

; APPLICANT: DUBELL, Arnold T

; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L

; APPLICANT: JIANG, Cai-Zhong

; APPLICANT: REUBER, T. Lynne

; APPLICANT: CREELMAN, Robert A

; APPLICANT: PINEDA, Omaira

; APPLICANT: YU, Guo-Liang

; APPLICANT: BROWN, Pierre E

; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants

; FILE REFERENCE: MEI0036-2 US

; CURRENT APPLICATION NUMBER: US/10/225,066A

; CURRENT FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 09/837,444

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-12-05

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 1122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 365

; LENGTH: 1818

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-10-225-066A-365

Query Match 18.7%; Score 377.4; DB 13; Length 1818;
 Best Local Similarity 76.7%; Pred. No. 1.5e-93;
 Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY 702 GCTACACCGAAGAAACTATTGAGAGTTTGGACAGAGACGCTCTATATACCGCGTGT 761
 Db 554 GCCACGCCAAGACGTGATTGGACATTTTCGGACAACGACCTCGATCTATCGTGTGTC 613
 QY 762 ACAAGGCATCGTGGACAGGAAGATATGAGGCACATTTATGGGATAATAGTTTGTAAAGA 821
 Db 614 ACAAGACATCGATGAGTGTGATATGAGGCTCATCTATGGGATAATAGTTTGTAAAGG 673
 QY 822 GAAGGCCAAACGCGCAAGGAAGACAAAGTTTATTTTCGGAGGTTATGACAAAAGAAAGAAA 881
 Db 674 GAAGGCCAGTCTTAGAAGGAAGACAAAGTTTATCTTGGTGGATATGACAAAGAGATAAA 733
 QY 882 GCAGCTAGGCTTTATGATTTAGCGCACTCAAGTATTTGGGAAACCACTACTACTAAAC 941
 Db 734 GCAGCAAGATCATATGATCTAGCTGCACCTTAGTACTGGGCTCCTTCAACTACTACTAAT 793
 QY 942 TTCCCCATGAGCGAATATGAAAAGAGGTAGAGAGATGAAGCACAATGACAAGGCAAGAG 1001
 Db 794 TTCCCCATTTCAAACCTACGAGAAAGAGTATAGGAATTAAGCAATGACGACGACAGAG 853
 QY 1002 TATGTTCCCTCACTGCGCAGGAAAAGTAGTGGTTTCTCTCGTGGTGCATCGATTTATCGT 1061
 Db 854 TTCGTGGCTGCCATTAGAAGGAAAAGTAGTGGATTTTCGAGAGGCGCTTCGATGTATCGA 913
 QY 1062 GGCTAACAGACATCACCAACATGGAAGATGCGAAGTAGGATAGGAAGAGTCGCCGCT 1121
 Db 914 GGATTTCAAAGGCATCACCAACATGGAAGATGCAAGCAAGAGTATCGCCGAGTCGCCGGA 973
 QY 1122 AACAAAGACCTTCACTTGGAACTTTTGGCACAACAAGAAAGAGTCGACAGGAGATACGAC 1181
 Db 974 AACAAAGACCTTACTTTGGAACTTTTAGCACCTGAGGAAGAGCAGCAGAGCTTACGAT 1033
 QY 1182 ATTGCGGCCATCAAAATTCAGAGGATTAACCGCAGTCACCTAACTTCGACATGAACAGATAC 1241
 Db 1034 ATAGTCGAATTAAGTTTAGGACCTTAATGCAAGTACCAACTTCGAGATCAACCGGTAC 1093
 QY 1242 AACGTTAAAGCAATTCCTGAAAGCCCTAGTCTTTCCTAT--TGGTAGCCCGCAAAACGT 1298
 Db 1094 GACGTGAAAGCCATTCTAGAGAGTAGCACTCTTCCCATCGGAGGAGCGCAGCTAAACGG 1153
 QY 1299 CTCGAAGGAGGCTAACCGTC 1317
 Db 1154 CTCGAAGAGCTCAAGCTC 1172

RESULT 8

US-10-225-066A-855

; Sequence 855, Application US/10225066A

; Publication No. US20030226173A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: RATCLIFFE, Oliver

; APPLICANT: RIECHMANN, Jose Luis

; APPLICANT: ADAM, Luc J

; APPLICANT: DUBELL, Arnold T

; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L

; APPLICANT: JIANG, Cai-Zhong

; APPLICANT: REUBER, T. Lynne

; APPLICANT: CREELMAN, Robert A

; APPLICANT: PINEDA, Omaira

; APPLICANT: YU, Guo-Liang

; APPLICANT: BROWN, Pierre E

; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants

; FILE REFERENCE: MEI0036-2 US

; CURRENT APPLICATION NUMBER: US/10/225,066A

; CURRENT FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 09/837,444

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-12-05
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 1122
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 855
 ; LENGTH: 1818
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-225-066A-855

Query Match 18.7%; Score 377.4; DB 13; Length 1818;
 Best Local Similarity 76.7%; Pred. No. 1.5e-93;
 Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY	702	GCTACACCGCAAGAACTATTGAGAGTTTGGACAGAGGAGCTCTATATACCGGGTGT 761
DB	554	GCACGCCAAGACGTGCTATGGACACTTTCGACAAACCACTCGATCTATCGTGGTGC 613
QY	762	ACAAGGCATCGTGACAGGAAGATATGAGGCACATTTATGGGATAATAGTTCTAAAGA 821
DB	614	ACAAGACATCGATGACTGGTGCATATGAGGCTCATCTATGGGATAATAGTTGTAGAAG 673
QY	822	GAAGCCAAACCGCAAGAGACAGTATTATTTGGAGGTTATGACAAAGAGAAAA 881
DB	674	GAAGGCCATCTAGAAAGAGAGACAGTTTACITGGGTGGATATGACAAAGAGATAA 733
QY	882	GCAGCTAGGGCTTATGATTTAGCCGCACTCAAGTATTTGGGAAACCACTACTACTAAC 941
DB	734	GCAGCAAGATCATATGATCTAGCTGCATTAAGTACTGGGTCTTCAACTACTACTAAT 793
QY	942	TTCCCATCAGCGAATATGAAAGAGGCTAGACAGATGACACATGACAGCAAG 1001
DB	794	TTCCCATTTACAACTACGAGAAAGATGAGGAAATGAAAGCACTGACAGACAAG 853
QY	1002	TATGTTGCCCTACTCGCAGGAAAAAGTAGTGTCTCTCTGTTGTCATCGATTATCGT 1061
DB	854	TTCTGTTGCTGCCATTAGAAAGAAAAAGTAGTGTCTCTCTGTTGTCATCGATTATCGA 913
QY	1062	GGAGTAAAGCACTCCACATCGAGGATTAACCGCAGTACTACTTTCGACATGACAGATAC 1121
DB	914	GGAGTTACAAAGTTTAGAGACTTAATGAGTGGACCACTTCGAGATCAACCGGTAC 1093
QY	1242	AACGTTAAAGCACTCCGAAAGCCCTAGTCTTCTCTAT---TGGTAGCGCGCAAAACGT 1298
DB	1094	GACGTGAAAGCCATTCTAGAGAGTAGCACTCTCTCCATCGGAGGAGCGGAGCTAAACGG 1153
QY	1299	CTCAAGGAGGCTAACCGTC 1317
DB	1154	CTCAAGAGAGCTCAAGTC 1172

RESULT 9
 US-10-225-066A-911
 ; Sequence 911, Application US/10225066A
 ; Publication No. US20030226173A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: RIECHMANN, Jose Luis
 ; APPLICANT: ADAM, Luc J
 ; APPLICANT: DUBELL, Arnold T
 ; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L
 ; APPLICANT: JIANG, Cai-Zhong
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: CREELMAN, Robert A
 ; APPLICANT: PINEDA, Omaira
 ; APPLICANT: YU, Guo-Liang
 ; APPLICANT: BROWN, Pierre E
 ; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: MB10036-2 US
 ; CURRENT APPLICATION NUMBER: US/10/225,066A
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 09/837,444
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/310,847
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/336,049
 ; PRIOR FILING DATE: 2001-12-05
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 1122
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 911
 ; LENGTH: 1818
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-225-066A-911

Query Match 18.7%; Score 377.4; DB 13; Length 1818;
 Best Local Similarity 76.7%; Pred. No. 1.5e-93;
 Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY	702	GCTACACCGAAGAACTATTGAGAGTTTGGACAGAGGAGCTCTATATACCGGGTGT 761
DB	554	GCACGCCAAGACGTGCTATGGACACTTTCGACAAACCACTCGATCTATCGTGGTGC 613
QY	762	ACAAGGCATCGTGACAGGAAGATATGAGGCACATTTATGGGATAATAGTTGTAAAGA 821
DB	614	ACAAGACATCGATGACTGGTGCATATGAGGCTCATCTATGGGATAATAGTTGTAGAAG 673
QY	822	GAAGCCAAACCGCAAGAGACAGTATTATTTGGAGGTTATGACAAAGAGAAAA 881
DB	674	GAAGGCCATCTAGAAAGAGAGACAGTTTACTTGGGTGGATATGACAAAGAGATAA 733
QY	882	GCAGCTAGGGCTTATGATTTAGCCGCACTCAAGTATTTGGGAAACCACTACTACTAAC 941
DB	734	GCAGCAAGATCATATGATCTAGCTGCATTAAGTACTGGGTCTTCAACTACTACTAAT 793
QY	942	TTCCCATCAGCGAATATGAAAGAGGCTAGACAGATGACACATGACAGCAAG 1001
DB	794	TTCCCATTTACAACTACGAGAAAGATGAGGAAATGAAAGCACTGACAGACAAG 853
QY	1002	TATGTTGCCCTACTCGCAGGAAAAAGTAGTGTCTCTCTGTTGTCATCGATTATCGT 1061
DB	854	TTCTGTTGCTGCCATTAGAAAGAAAAAGTAGTGTCTCTCTGTTGTCATCGATTATCGA 913
QY	1062	GGAGTAAAGCACTCCACATGAGGATTAACCGCAGTACTACTTTCGACATGACAGATAC 1121
DB	914	GGAGTTACAAAGTTTAGAGACTTAATGAGTGGACCACTTCGAGATCAACCGGTAC 1093
QY	1122	AACAAAGACCTCTACTTGGGAACTTTTGGCACACAGAGAGAGCTCCAGAGGATACGAC 1181
DB	974	AACAAAGACCTCTACTTGGGAACTTTTAGCACTGAGGAAAGAGAGAGCTCCGCGGAG 1033
QY	1182	ATTGGGCGCATCAAAATTCAGAGGATTAACCGCAGTACTACTTTCGACATGACAGATAC 1241
DB	1034	ATAGCTGCATTAAGTTTAGAGACTTAATGAGTGGACCACTTCGAGATCAACCGGTAC 1093
QY	1242	AACGTTAAAGCACTCCGAAAGCCCTAGTCTTCTCTAT---TGGTAGCGCGCAAAACGT 1298
DB	1094	GACGTGAAAGCCATTCTAGAGAGTAGCACTCTTCCATCGGAGGAGCGGAGCTAAACGG 1153

QY 1299 CTCAAGGAGGCTAACCGTC 1317
DB 1154 CTCAAGAGAGCTCAAGCTC 1172

RESULT 10

US-10-374-780A-333

; Sequence 333, Application US/10374780A

; Publication No. US2004001927A1

; GENERAL INFORMATION:

; APPLICANT: Sherman, Bradley K

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Heard, Jacqueline E

; APPLICANT: Haake, Volker

; APPLICANT: Creelman, Robert A

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Adam, Luc J

; APPLICANT: Reuber, T. Lynne

; APPLICANT: Keddie, James

; APPLICANT: Brown, Pierre E

; APPLICANT: Pilgrim, Marsha L

; APPLICANT: Dubell III, Arnold T

; APPLICANT: Pineda, Omaira

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

; FILE REFERENCE: MBI-0047 CIP

; CURRENT APPLICATION NUMBER: US/10/374,780A

; CURRENT FILING DATE: 2003-02-25

; PRIOR APPLICATION NUMBER: 09/837,944

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 09/934,455

; PRIOR FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 10/225,066

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,067

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,068

; PRIOR FILING DATE: 2002-08-09

; NUMBER OF SEQ ID NOS: 2906

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 333

; LENGTH: 1818

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G1793

US-10-374-780A-333

Query Match 18.7%; Score 377.4; DB 16; Length 1818;

Best Local Similarity 76.7%; Pred. No. 1.5e-93;

Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY 702 GCTACACCGAAGAAAACCTATTGAGAGTTTGGACAGAGGACGCTCTATATACCGCGGTGT 761

DB 554 GCCACGCCAAGACGTGCTGAGACCTTTCGACAAAGAACCTCGATCTATCGTGTGTC 613

QY 762 ACAAGGATCCGTCGACAGGAGATATGAGGCACATTATCGGATATATAGTTGTAAGA 821

DB 614 ACAAGATCATGAGCTGCTCGATATGAGGCTCACTATGGATATAGTTGTAGAAG 673

QY 822 GAAGCCCAAGCCGCAAGGAAGCAAGTTTATTGGAGGTTATGACAAAGAGAAAAA 861

DB 674 GAAGCCGAGTCTAGGAAGGAGAACAGTTTACTTTGGTGGATATGACAAAGAGATAA 733

QY 882 GCAGTAGGGCTTATGATTATTAGCCGACTCAAGTATTGGGGAACCACTACTACTAAC 941
DB 734 GCAGTAGATCATATGATCTAGCTGCACCTAAGTACTGGGTCCTTCAACTACTACTAAT 793
QY 942 TTCCCATGAGCGAATATGAAGAAGAGGTAGAGAGATGAAGCATGACAGGCGAAG 1001
DB 794 TTCCCATTTACAAACACGAGAAAGAGTAGAGAAATGAAGCATGACAGGAGAGAG 853
QY 1002 TATGTTGCCTCACTCGCCGAGGAAAAGTAGTGGTTTTCTCTCGTGGTGCATCGATTTATCGT 1061
DB 854 TTGTTGCTGCCATTAGAAGGAAAAGTAGTGGATTTCGAGAGGCGCTTCGATGATCGA 913
QY 1062 GGAGTACAGAGATCACCAACATGGAAGATGGCAAGTGGCAAGTAGTAGGAGAGTCCCGGT 1121
DB 914 GGAGTTACAAGGATCACCAACATGGAAGATGGAAGATGGAAGGAGTCCCGGATCGCCGGA 973
QY 1122 AACAAAGACCTCTACTTTGGGAACCTTTTGGCACACAAGAAGAAGCTGCAGAGGCATACGAC 1181
DB 974 AACAAAGACCTCTACTTTGGGAACCTTTTGGCACTGAGGAAGAAGCAGCAGAAGCTTACGAT 1033
QY 1182 ATTGCGGCCATCAATTCAGAGGATTAACCGCAGTGACTTAATTCGACATGAACAGATAC 1241
DB 1034 ATAGCTGCAATAAAGTTTAGAGGACTTAATGCAAGTGAACCACTTCGAGATCAACCGGTAC 1093
QY 1242 AACGTTAAGCAATCCCTCGAAAGCCCTAGTCTTCTCTAT---TGGTAGCCGCCGCAAAAGCT 1298
DB 1094 GAGTGAAGGCCATTCAGAGAGTAGCACTCTTCCCATCGGAGGAGGCGCAGCTAAACGG 1153
QY 1299 CTCAGAGGAGCTAACCGTC 1317
DB 1154 CTCAGAGAGCTCAAGCTC 1172

RESULT 11

US-10-425-114-20762

; Sequence 20762, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 20762

; LENGTH: 794

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3245-185-P4_FLI

US-10-425-114-20762

Query Match 18.1%; Score 364.4; DB 13; Length 794;

Best Local Similarity 76.1%; Pred. No. 3.4e-90;

Matches 449; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 720 ATTGAGAGTTTGGACAGAGGAGCTCTATATACCGCGGTGTTACAGGATCGGTGACA 779

DB 1 ATCGACACGTTCCGGCAAGGACCTCTATATATCGAGGTGTAAACAGGATAGATGACA 60

QY 780 GGAAGATATCAGGCACATTTATGGGATAATAGTTGTAAAGAGAGAGGCCAAACGCGCAA 839

DB 61 GGGCGGTATGAGGCTCATCTATGCGGATATAGTTGTAGAGGAGGAGGAGGAGTGGCAG 120

QY 840 GGAAGCAAGTTTATTGGAGGTTATGACAAAGAGAAAAAGCAGCTAGGCTTATGAT 899

DB 121 GGTAGGCAAGTTTACCTTGGTGGTATGACAGGAGGAGCAAGGAGGAGGAGGCTTACGAT 180

APPLICANT: Cac, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 14859
LENGTH: 1479
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: LIB3029-010-B7_FLI
US-10-425-114-14859

Query Match 17.7%; Score 357.4; DB 13; Length 1935;
Best Local Similarity 73.8%; Pred. No. 5.e-88;
Matches 454; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 658 ACAACGCTTGGCCCAAGGAAGACTATTGATGATAGCGTTGAAGCTACACGGAAGAAA 717
Db 442 AATCGATATTCACAAACACAGCAAGCACTATAGCTGAGCGCTGCACCCAAAAGAT 501
QY 718 CTATTGAGAGTTTGGACAGAGAGCTCTATATACCGCGTCTTACAGGCTACGTTGGA 777
Db 502 CTGTTGATATCTTTGGCCCAAGCTACCTCGATCTACCGGGGTGTACCAGGCATAGTGA 561
QY 778 CAGGAAGATATGAGGCACATTTATGGGTAATAGTTTAAAGAGAGAGCCCAAGCCGCA 837
Db 562 CTGGAAGGTATGAGCTCACTTGTGGGCAATAGTTGCAAGAGAGAGCCCAAGTGA 621
QY 838 AAGGAAGCAAGTTTATTTGGGAGTTTATGACAAAGAGAGAGAGAGAGAGAGAGAG 897
Db 622 AAGGAAGCAAGTTTATTTGGGTGGTATGACAAAGAGAGATAAAGCTGCAAGAGCTTATG 681
QY 898 ATTTAGCCCACTCAAGTATTTGGGAAACCACTACTACTACTTCCCATGAGCGAAT 957
Db 682 ATCTTGGGCTCTCAAGTATCTGGGTCCGACCACTACTACTACTTCCGATTTCCAACT 741
QY 958 ATGAAAAAGAGTTAGAGAGATGAAGCAATGACAAAGCAAGAGTATGTTGCTCTACTGC 1017
Db 742 ATGAGAGAGAGCTCGAAGAGATGAAGCAATGACTAGGCAAGAGTTCTGTTGCTCTCTCC 801
QY 1018 GCAGGAAGAGTATGTTTCTCTGTTGCTGATCTGATTTATCGTGGATGAAGAGATC 1077
Db 802 GAGGAAGAGTATGTTTCTCTGTTGCTGATCTGATTTATCGTGGATGAAGAGATC 861
QY 1078 ACCAATGGAAGATGCAAGCTAGGATAGGAAGAGTTCGCGGTAAACAAGAGCTCTACT 1137
Db 862 ATCAACATGTTAGTGGCAAGCAAGATTCGAAGAGTTGCAAGCAACAAGATCTCTATC 921
QY 1138 TGGGAATTTTGGACACAAAGAGAGCTGCAAGGATAGACATGCGGCCATCAAT 1197
Db 922 TTGSCACATTTAGCACCAAGAGAGAGAGCTGAAGCCTATGATATTGCAAGCAATCAAGT 981
QY 1198 TCAGAGGATTAACCGCAGTCACTAACTTCGACATGAACAGATCAACGTTTAAAGCAATCC 1257
Db 982 TTAGAGGCTTAACCGCGTTACAAATTTGATATGAGCCCTACGATGTAACAGCATTTG 1041
QY 1258 TCGAAAGCCCTAGTC 1272
Db 1042 CAAACAGCAATCTTC 1056

RESULT 14
US-10-425-114-14859
Sequence 14859, Application US/10425114
Publication No. US20040034889A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kowalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.

APPLICANT: Cac, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 14859
LENGTH: 1479
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: LIB3029-010-B7_FLI
US-10-425-114-14859

Query Match 17.7%; Score 357; DB 13; Length 1479;
Best Local Similarity 76.4%; Pred. No. 6.e-88;
Matches 438; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 700 AAGTACACCGAAGAAAACCTATTGAGAGTTTGGACAGAGAGCTCTATATACCGCGGTG 759
Db 38 AACCTCTCCAAAGAAAACCGTCGACACCTTCGGCCAAACGACCTCCATCTACCGCGCG 97
QY 760 TTACAAGGCATCGGTGGACAGAGATATGAGGCACATTTATGGGTAATAGTTGTAAAA 819
Db 98 TCACCCACATAGATGACGGAAGATACGAAGCTCATCTATGGGCAATAGTTGTAGAA 157
QY 820 GAGAAGGCCAAAACCGCAAGAGAAAGCAAGTTTATTTGGGAGGTATGACAAAGAGAA 879
Db 158 GAGAAGGCCAAAACCGCAAGAGAAAGCAAGTTTACCTGGGTGTTATGACAAAGAGATA 217
QY 880 AAGCAGCTAGGCTTATGATTTAGCCGACCTCAAGTATTTGGGAAACCACTACTACTA 939
Db 218 AGGAGAGCCAGGCTTACGATCTCGAGCTCTCAAGTACTGGGTCCAACCTACCACCA 277
QY 940 ACTTCCCATGAGGCAATATGAAAAAGAGGTAGAAAGATGAAGCAATGACAAAGCAAG 999
Db 278 ACTTCCCATTTCCAACTATGAGAAGGAACTGGAGAGATGAAGCAATGACAAAGCAAG 337
QY 1000 AGTATGTTGCTCTACTGCGAGGAAAGTATGTTTCTCTCGTGGTGCATCGATTTATC 1059
Db 338 AGTTGTTGCTCTACTGAGAGAGAGAGTATGTTTCTTAGGGGGCTCTATATACA 397
QY 1060 GTGAGTAAACAGACATCAACCAAGATGAAGATGGCAAGCTAGGATAGGAAGTTCGCG 1119
Db 398 GAGGAGTCAGGAGACACCAACGAGATGCGCATGCGAGGAGATAGGAGATGCGCG 457
QY 1120 GTACAAAGAGCTCTACTTGGAACTTTGGACACAGAGAGAGCTGACAGAGCATACG 1179
Db 458 GAAACAAAGAGCTCTACTTGGAACTTTGACGACCAAGAGAGAGCTGCTGAGGCTATG 517
QY 1180 ACATTGCGGCCATCAAAATTCAGAGGATTAACCGAGTCACTAACTTCGACATGAACAGAT 1239
Db 518 ACATTGCTGCTATCAAAATTCAGGAGATTAATGCAATGCAAACTTTGACATGAGTCGCT 577
QY 1240 ACAAGCTTAAAGCAATTCCTCGAAAGCCCTAGTC 1272
Db 578 ACCATGTAAGAGAGCATTCGAAATAGCACTCTTC 610

RESULT 15
US-10-424-599-70525
Sequence 70525, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kowalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 14:52:42 ; Search time 4152 Seconds
(without alignments)
4164.307 Million cell updates/sec

Title: US-09-980-364-2
Perfect score: 3119
Sequence: 1 MNNWLGSLSPYEQNHHRK.....YHGSGGGEVAPFTVWNND 579

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USP10.spool/US09980364/runat_15072004_095236_25880/app_query.fasta_1.775
-DB=EST -Qfmt=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09980364@cgn 1.1 3437 @runat_15072004_095236_25880 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELENT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	992	31.8	769	14	CA783156	
2	949.5	30.4	1753	11	AY109146	
C	3	947.5	30.4	556	28	BZ512981
	4	940.5	30.2	882	14	CK267021
	5	934.5	30.0	590	12	BM094116
	6	899.5	28.5	748	14	CF807326
	7	875	28.1	720	13	BF864461
	8	872.5	28.0	552	13	BQ864451
	9	869.5	27.9	593	12	BM307997
10	851.5	27.3	644	10	AW200688	
11	849	27.2	1160	14	CK206573	
12	841	27.0	555	12	BJ188928	
13	841	27.0	558	12	BJ178045	
14	838	26.9	690	13	CA094556	
15	834	26.7	500	10	AW780688	
16	820	26.3	545	14	CD475882	
17	815	26.1	631	13	CA103041	
18	811	26.0	674	10	BF647766	
19	809	25.9	558	14	CA232734	
20	809	25.9	599	14	CA230632	
21	796	25.5	697	14	CA189006	
22	781.5	25.1	585	12	BM086088	
23	781.5	25.1	640	13	CA100375	
24	743.5	23.8	641	14	CD879292	
25	713	22.9	420	9	AJ475492	
26	700.5	22.5	466	13	BU965263	
27	699	22.4	596	12	BM307877	
28	699	22.4	935	14	CK261980	
29	694	22.3	704	13	BQ625052	
30	691	22.2	907	14	CK289457	
31	688.5	22.1	469	14	CF324329	
C	32	688	22.1	497	14	CF244784
	33	685.5	22.0	842	14	CF446573
	34	676.5	21.7	513	13	BQ123372
	35	675.5	21.7	545	12	BI974354
	36	672	21.5	1897	11	AY103852
	37	670.5	21.5	476	12	BM092890
	38	665.5	21.3	956	14	CK294077
39	655.5	21.0	955	14	CK257531	
40	645.5	20.7	948	14	CK259208	
41	645	20.7	679	12	BJ312281	
C	42	640.5	20.5	728	14	CF035870
	43	635.5	20.4	694	14	CF036846
	44	635.5	20.4	737	14	CD832004
	45	631	20.2	527	12	BM309412
	46	611	20.0	412	12	BM309412

ALIGNMENTS

RESULT 1
CA783156
LOCUS
DEFINITION
CA783156 769 bp mRNA linear EST 04-DEC-2002
sat20d05.v1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl036-14361 5' similar to TR:Q9SYC2 Q9SYC2 F1LM15.6 PROTEIN. ;
mRNA sequence.
CA783156
CA783156.1 GI:26045627
EST.
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE AUTHORS

1 (bases 1 to 769)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, F., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

TITLE JOURNAL COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: cu@resgen.com web site: www.resgen.com

Seq primer: -40RP from Gibco
High quality sequence stop: 454.
Location/Qualifiers

FEATURES source

1..769
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl036-14361"
/tissue_type="somatic embryos cultured on MSD 20"
/lab_host="DH10B"
/clone_lib="Gm-cl036"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Alignment Scores:

Pred. No.: 8,04e-87 Length: 769
Score: 992.00 Matches: 187
Percent Similarity: 89.21% Conservative: 28
Best Local Similarity: 77.59% Mismatches: 14
Query Match: 31.81% Indels: 12
DB: 14 Gaps: 3

US-09-980-364-2 (1-579) x CA783156 (1-769)

QY 160 LysGlyLeuSerLeuSerMetAsnSer-----SerThrSerCysAspAsnAsnAsp 177
Db 20 AATCATGACATATATCCATGGAGTGGTAAGATTCACATGTGAACACCGATGTGAA 79
QY 178 SerAsnAsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrPro 197
Db 80 AATAGCACAAACACT-----ACTGTCGAATAGCACCT 112
QY 198 LysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHis 217
Db 113 AGAAGAACTTGGATATCTGGCGAGAGAACATCATATATCGTGAGTAACATGACAT 172
QY 218 ArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGlyGln 237

Db 173 AGATGACTGGAAGGTATCAAGCTCATCTTTGGGATATAGCTGTAGAGGAGGCCAA 232
QY 238 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAlaArg 257
Db 233 TCAAGAAAAGGAGCGCAAGTTTATTGGGTGATATGATATAAAGAGAGAACACCTAGA 292
QY 258 AlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyTyrThrThrThrThrAsnPheProMet 277
Db 293 GCATTATGATTAGTCTCTGAAGTACTGGGGGACATCCACACTTACCACTTCCAATT 352
QY 278 SerGluTyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValala 297
Db 353 AGCAACTATTCAGAAGGAATGGATGAAATGAACACATGACGAGACAAGAATTTTGTGCC 412
QY 298 SerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 317
Db 413 GCCATTAGAGGAAAAGCAGTGGTTTCTCCAGGGTGCATCAATGATTCGTGGAGTTACA 472
QY 318 ArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAsp 337
Db 473 AGGCATCACCAACACGGAAGTGGCAAGCAAGGATTGGCAGAGTTGCAGGAAACAAGAT 532
QY 338 LeuTyrLeuGlyThrPheGlyThrGlnGluAlaGluAlaTyrAspIleAla 357
Db 533 CTTTACTTGGGAACCTTCAGTACTGAGGAAGAGGCTGCAGAGGATACGACATAGCAGCG 592
QY 358 IleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLys 377
Db 593 ATAAAGTTTCAAGAGTCTCAACGCTGTCAAACTTTTGACATGAGCCGATACGACGTGAA 652
QY 378 AlaIleLeuGluSerProSerLeuProIle---GlySerAlaAlaLysArgLeuLysGlu 396
Db 653 GCATTCCTTGAAGCAACACTCTCCCAATAGGAGGCGCTGCAAAAGCGTCTGAAAGAA 712
QY 397 Ala 397
Db 713 GCT 715

RESULT 2

LOCUS AY109146 1753 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays PCO137288 mRNA sequence.
ACCESSION AY109146
VERSION AY109146.1 GI:21212582
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1753)
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1753)
AUTHORS Coe, E.H.

TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSI, maizeMap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES source

Location/Qualifiers
1..1753
/organism="Zea mays"
/mol_type="mRNA"

```

/db_xref="MaizeDB:638053"
/db_xref="taxon:4577"
/cloned_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.: 4,6e-82 Length: 1753
Score: 949.50 Matches: 233
Percent Similarity: 48.46% Conservative: 50
Best Local Similarity: 39.90% Mismatches: 130
Query Match: 30.44% Indels: 171
DB: 11 Gaps: 18

US-09-980-364-2 (1-579) x AY109146 (1-1753)
QY 3 AsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArg-----Lys 20
DB 63 CACCATTGGCTCTCCTCTCCTCTCCTC-----AACCACTACCACCATGGCTACTC 113
QY 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
DB 114 GAAGCCTTCTAACTCTCCGTACTCTCTT-----146
QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
DB 147 -----GGAGACGACGAGCGGCGCACTGGAGGATCC-----176
QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsn 80
DB 176 -----176
QY 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLys---LeuGluAsn 99
DB 177 -----CCGAGGACGGTGGAGGAC 194
QY 100 PheLeuGlyArgThrThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
DB 195 TTCTTCCGCG-----GGC 206
QY 120 SerGlyCysTyrGlyGlyAspGlyGly-----130
DB 207 GTCCGTTGGTCGGCGCCCGCCGCGCGCTGCAGATCAGGATCACCAGCTGTG 266
QY 131 ---GlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsn 149
DB 267 TCGCGCGAGCTGGGC---AGCATCAGACGCCAGGTCTTTCGCGCCACTACCCGCGCGGCCCA 323
QY 150 ValAspAsnGln---GluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSer 168
DB 324 GCTGGGACGACGCTGGAGAACCCCGCGCGGTGACCGTGGCGGCATGCTGTCGACGGAC 383
QY 169 SerThrSerCysAspAsnAsnAspSerAsnAsnAsnValValAlaGlnGlyLysThr 188
DB 384 GTGCGGGGGCGGAGTCCGACCGAG-----407
QY 189 IleAspAspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThr 208
DB 408 -----GCGAGGCGGCGCCGCGAGAGCTTCGCGCCAGCGACA 443
QY 209 SerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrp 228
DB 444 TCATCTACCTACCGTGGCGTCCACGAGGACCGGTGGACGGGAGATATGAGGCGCACCTGTGG 503
QY 229 AspAsnSerCysLysArgGluGlnThrArgLysGlyArgGlnValTyrLeuGlyGly 248
DB 504 GACACACACTGCGCGCGGAGGGCCAAAGCGCAAGGACGCCCAAGTCTACCTAGGAGGC 563

```

RESULT 3

BZ512981/c

LOCUS

DEFINITION

BZ512981

556 bp DNA linear

GSS 16-DEC-2002

BOMQP05TF BO_2_3_KB Brassica oleracea genomic clone BOMQP05,

```

QY 249 TyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGly 268
DB 564 TATGACAGAGGAGAGAGAGCGGCTAGAGCTTACGACCTCCCGCGCTCAAGTACTGGGG 623
QY 269 ThrThrThrThrAsnProMetSerGluTyrGluLysGluValGluGluMetLys 288
DB 624 CCTACACACGACCACTTCCAGGTGTCCAACTTACGAGAAGGAGCTGGAGGAGATGAAA 683
QY 289 HisMetThrArgGlnGluTyrValAlaSerLeuArgLysSerSerGlyPheSerArg 308
DB 684 TCCATGACCGCGAGGAGTTCATCGCTGTGGCAGAGAGCAGCGCTTCTACGA 743
QY 309 GlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGlnAlaArg 328
DB 744 GCGCGCTCATCTACAGAGGAGTCAAGGCATCATCAGCAGCGCGCTGCGAGGAGG 803
QY 329 IleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGlu 348
DB 804 ATCGGAGGTTGGCGGGAACAGGACCTGTACTTGGGCACCTTTCAGTACTCAGGAAGAG 863
QY 349 AlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsn 368
DB 864 GCGGCGGAGGCTACGACATCGCTGCATCAAGTTCGCGCGGCTCAACGCGCTCACCAC 923
QY 369 PheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuProIleGly 388
DB 924 TTTGATCATCGCGCTAGCAGCTGGAGGATCTCTCAGCAGC---GACCTCCCGCTCGG 980
QY 389 SerAlaAlaLysArgLeuLysGluAlaAsnArgProValProSerMet-----404
DB 981 GCGGAGTAGCGGTGCGCGCGCGCCCAAGTTCGCTGGACTCGCTGACGCGGGGAGC 1040
QY 405 -----MetMetIleSerAsnAsnValSerGluSerGlu-----Asn 416
DB 1041 GCTCGCGCATGATCTCGCGCGGCTGCTGCGCTTCGAGCGCCACCATTGCGCGCTCC 1100
QY 417 SerAlaSerGlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeu 436
DB 1101 GAGAGGACTACTGCTCTGCTCGCGCTGCACTTACCAGCAG-----1142
QY 437 LeuHisGlnHisGlnGluArg-----TyrAsnGlyTyrTyrTyrAsn 450
DB 1143 ---CAGCAGGAGGAGGAGGAGGAGTTCGCGCTTCGAGCGCCACCATTGCGCGCTCC 1199
QY 451 GlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheLysGlnGluAspAspGlnHis 470
DB 1200 GCGTGAACGTGGAC-----1214
QY 471 HisPheLeuSerAsnThrGlnSerLeuMetThrAsn-----482
DB 1215 ---TTCAGATGGGACCCAGCGCGGAGCAACAAACACCGGACGCGCGCTCATGTGG 1271
QY 483 -----IleAspHisGlnSerSerValSerAspAspSerValThr 495
DB 1272 GCGCGCACCTCTGTGTGAGTAGTGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1331
QY 496 ValCysGlyAsn-----ValValGlyTyrGlyGlyTyrGlnGly 508
DB 1332 TATGCCAGTAACTCTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1391
QY 509 PheAlaAlaProValAsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArg 528
DB 1392 TCCACCGCGCAATGGAACTGGTGTACTACGACTACGAGCAACACCGGACCGCGCT 1451
QY 529 AsnHisTyrTyr 532
DB 1452 CCCCACTACTAC 1463

```

genomic survey sequence.
 BZ512981
 BZ512981.1 GI:27039560
 GSS.
 ORGANISM
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 556)
 Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished (2001)
 Other GSSs: BOMQPO5TR
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.

FEATURES
 source
 1..556
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="Tol000DH3"
 /db_xref="taxon:3712"
 /clone="BOMQPO5"
 /clone_lib="BO_2_3 KB"
 /note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared
 Genomic DNA inserted into pHO51 using BstXI linkers"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.14e-82 Length: 556
 Score: 947.50 Matches: 183
 Percent Similarity: 98.39% Conservative: 0
 Best local Similarity: 98.39% Mismatches: 3
 Query Match: 30.38% Indels: 2
 DB: 28 Gaps: 0

US-09-980-364-2 (1-579) x BZ512981 (1-556)

QY 377 LysalalleuGluSerProSerLeuProGlySerAlaAlaLysArgLeuLysGlu 396
 DB 556 AAGCATCCCTCGAAGCCCTAGTCT-CCTATTGT-AGCGCGCAAAACGTTCAAGGAG 499
 QY 397 AlaAsnArgProValProSerMetMetMetIleSerAsnValSerGluSerGluAsn 416
 DB 498 GCTAACCGTCGGTTCGAAGTATGATGATCATGATAGTAAACGTTTCAGAGAGTGAGAT 439
 QY 417 SerAlaSerGlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeu 436
 DB 438 AGTGCTAGCGGTTCGCAAAACGTCGCGTTCAGATCATCAGGAGTAGATTGACCTTA 379
 QY 437 LeuHisGlnHisGlnGluArgTyrAsnGlyTyrTrpAsnGlyGlyAsnLeuSerSer 456
 DB 378 TTGCACCAACATCAAGAGAGTCAATGGTTATATTACAAATCGAGGAAACTTGTCTTCG 319
 QY 457 GluSerAlaArgAlaCysPheLysGlnGluAspAspGlnHisPheLeuSerAsnThr 476
 DB 318 GAGAGTCTAGGCTTCTTTCAACAGAGAGATGATCAACACCATTTCTTGACCAACAG 259
 QY 477 GlnSerLeuMetThrAsnIleAspHisGlnSerSerValSerAspSerValThrVal 496
 DB 258 CAGAGCCTCATGACTATATATCGATCATCAAGTCTCTGTTTCGGATGATTCGGTTACTGT 199
 QY 497 CysGlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAsp 516
 DB 198 TGTGGAAATGTTGTTGTTATGTGGTTATCAAGGATTTGCAGCCCCGGTTAACTCCGAT 139

QY 517 AlaTyrAlaAlaSerGluPheAspTyrAsnAlaAArgAsnHisTyrTyrPheAlaGlnGln 536
 DB 138 GCCTACGCTGCTAGTGTGATTATTAACGCAAGAAACCATTTACATTGCTCAGCAG 79
 QY 537 GlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMetThrAsnValGly 556
 DB 78 CAGCAGACCAGCAGTCGCCAGGTGGAGATTTTCCCGCGCAATGACGAATAATGTTGC 19
 QY 557 SerAsnMetTyrTyrHis 562
 DB 18 TCTAATATGTTATTACCAT 1

RESULT 4
 CK267021
 LOCUS
 DEFINITION
 EST713099 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAC822 5' end, mRNA sequence.
 CK267021
 CK267021.1 GI:39823999
 EST.
 Solanum tuberosum (potato)
 ORGANISM
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 882)
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST713100
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source
 1..882
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAC822"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1 d,
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.17e-81 Length: 882

RESULT 5

QY 223 TyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArg 242
 Db 61 TATGAAGCTCACCTTTGGGATAATAGCTGTGAAGGGAAGGCAATCAAGAAAGGAGCG 120
 QY 243 GlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAla 262
 Db 121 CAAAGTTATTTGGTGGATATGATAAAGAAAGAAAGAGCTAGTCTTATGATTAGCT 180
 QY 263 AlaLeuLysTyrTrpGlyThrThrThrThrThrThrThrThrThrThrThrThrThr 282
 Db 181 GCACCTCAAGTACTGGGGACATCCACCACTACCACTTTCCATTAGTAATCATGAGAG 240
 QY 283 GluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgLys 302
 Db 241 GAATTCGATGAATGAACACATGACGCGACAGAAATTTGTTGCTGCCATTAGAAGAAA 300
 QY 303 SerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHis 322
 Db 301 AGCAGTGGTTTCTCCAGGGGTGCATCAATGATATCTGAGTTACAGGCATCACCAAC 360
 QY 323 GlyArgTrpGluAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThr 342
 Db 361 GGAAGATGGCAAGCAAGAAATTTGCGAGAGTTGCGAGAAACAAAGATCTTTACTTGGAACT 420
 QY 343 PheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGly 362
 Db 421 TTCAGTACTGAGAGAGAGGCTGCTGAGCAGTACGATAGTGCATTAAGTTCAGAGGT 480
 QY 363 LeuThrAlaValThrAspPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSer 382
 Db 481 CTCAACGGTGTCAACAACTTTGACATGAGCGCTACGACGTGAAAGCCATCCTTGAAGC 540
 QY 383 ProSerLeuProIle---GlySerAlaAlaLysArgLeuLysGluAla 397
 Db 541 AACACTCTCCCAATAGGAGGAGGAGCTGCAAAAGCGTCTGAAGAAGCT 588

RESULT 6
 LOCUS CF807326 748 bp mRNA linear EST 27-OCT-2003
 DEFINITION psHB023xFO7f USDA-IFAFS:Expression of Phytophthora sojae genes during infection and propagation Phytophthora sojae cDNA clone SH023FO7 5, mRNA sequence.

ACCESSION CF807326
 VERSION CF807326.1 GI:37995737
 KEYWORDS EST.
 SOURCE Phytophthora sojae
 ORGANISM Phytophthora sojae
 Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.
 1 (bases 1 to 748)
 REFERENCE
 AUTHORS Tyler,B.
 TITLE Tyler,B. Not Published
 JOURNAL Unpublished (2003)
 COMMENT Contact: Tyler B
 Tyler lab
 VBI

1880 Pratt Dr., Blacksburg, VA 24061, USA
 Tel: 540-231-7318
 Email: bmtyley@vt.edu
 PCR Primers
 FORWARD: BK reverse primer
 BACKWARD: BK reverse primer
 Plate: 023 row: F column: 07
 Seq primer: BK reverse primer
 High quality sequence stop: 748.
 Location/Qualifiers
 1..748
 /organism="Phytophthora sojae"
 /mol_type="mRNA"
 /db_xref="taxon:67593"
 /clone="SH023FO7"
 /tissue_type="mycelium"

/cell_line="P6497"
 /dev_stage="48 hr. post infection stage"
 /lab_host="Soybean plant"
 /clone_lib="USDA-IFAFS:Expression of Phytophthora sojae genes during infection and propagation"
 /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Alignment Scores:
 Pred No.: 9,31e-77 Length: 748
 Score: 889.50 Matches: 174
 Percent Similarity: 81.51% Conservative: 20
 Best Local Similarity: 73.11% Mismatches: 19
 Query Match: 28.52% Indels: 25
 DB: 14 Gaps: 4
 US-09-980-364-2 (1-579) x CF807326 (1-748)
 QY 161 GlyLeuSerLeuSerMetAsnSerSerThrSerCys----- 172
 Db 39 GCGCGTTCTCCGCAACAGCCACGCGCACCGTGTGCACCGCAGCTTCCGAGTTC 98
 QY 173 -----AspAsnAsnAsn-----AspSerAsn----- 179
 Db 99 TCCACCGCACCAACCAACCCACTGTACGATTCCGAGCTGAAGACAAATAGCCCGGTGC 158
 QY 180 -----AsnAsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrPro 197
 Db 159 TTCCTCTCGCGCTTTGTGCGCGAACCACCAACC-----GAACCTCAGAAACCTCTCCA 212
 QY 198 LysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHis 217
 Db 213 AAAAAACCCGTGCACACCTTCGCCCAACGACCTCCATCTACCGCGGCTCACCAGAT 272
 QY 218 A:G:TrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGlyGln 237
 Db 273 AGATGGACGGGAAGATACGAAGCTCATCTATGGGACATAGTTGTAGAGAGAGGCCAA 332
 QY 238 ThrArgLysGlyArgGlnValTyrLeuGlyTyrAspLysGluGluLysAlaAlaArg 257
 Db 333 ACAGAGAAAGGAAGCAAGATTTCCTGGTGGTTATGACAAGGAAGATAAGCGAGCCAGG 392
 QY 258 AlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrThrThrThrThrThr 277
 Db 393 GCTTACGATCTCGAGCTCTCAAGTACTGGGTCCAACTACCACTACCACTTTCCCAAT 452
 QY 278 SerGluTyrGluLysGluValGluMetLysHisMetThrArgGlnGluTyrValAla 297
 Db 453 TCCAACTATGAGAGAACTGGAGGAGATGAAGAACATGACAGCAGGACAGAGTTTGTGCT 512
 QY 298 SerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 317
 Db 513 TCTCTACGAAGGAAGAGCGTGGTTTCTTAGGGGGCTCTATATACAGAGGAGTACG 572
 QY 318 ArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAsp 337
 Db 573 AGACACCAACGATGCGCGATGGCAGCGAGAATAGCAGAGTTGCCGGAACAAAGAC 632
 QY 338 LeuTyrLeuGlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAla 357
 Db 633 CTCTACCTTGGAACTTTTCAGCAACCAAGAAAGCTGCTGAGGCTATGACATTGTGCT 692
 QY 358 IleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsn 375
 Db 693 ATCANATTCAAGGATTAATGCACTGAGTCANCAACTTTTGACATGCTCCTCAGAT 746

FEATURES

source
 RESULT 7
 BQ864461
 LOCUS BQ864461
 DEFINITION QG26M12.yg.ab1 QG ABCDI lettuce salinas Lactuca sativa cDNA clone QG26M12, mRNA sequence.
 ACCESSION BQ864461
 VERSION BQ864461.1 GI:22249926

```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
  Koziak, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
  Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
  Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
  Lai, Z., Church, S., Jackson, L. and Bradford, K.
  Lettuce and Sunflower ESTs from the Compositae Genome Project
  http://compenomics.ucdavis.edu/
  Unpublished (2002)
  Contact: Alexander Koziak [R.W.Michelmore]
  Department of Vegetable Crops, R.W.Michelmore Lab
  University of California at Davis (UCD)
  Amundson Hall, UCD, Davis, CA 95616, USA
  Tel: 1-(530)-742-1742
  Fax: 1-(530)-752-9659
  Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
  singleton, see http://cgdb.ucdavis.edu/ for details.
  Plate: QGC26 row: M column: 12.

FEATURES
  source
    1..720
      /organism="Lactuca sativa"
      /mol_type="mRNA"
      /cultiivar="Salinas"
      /db_xref="taxon:4236"
      /clone="QGC26M12"
      /lab_host="E.coli"
      /clone_lib="QGC-ABCDI lettuce salinas"
      /notes="Vector: pERCNDSf1AB; The library was constructed
      from 10 different sources of RNA from a single genotype.
      Separate cDNAs were generated using primers that
      incorporated unique 5' and 3' tags to distinguish each
      source of RNA. cDNAs were then pooled, size-fractionated,
      directionally cloned into a custom medium-copy vector and
      transformationally made with four size classes to minimize
      size bias. Details of each source of RNA and library
      construction can be obtained at http://cgdb.ucdavis.edu/
      TAG_TISSUE=chemical induction
      TAG_LIB=QGC-ABCDI lettuce salinas
      TAG_SEQ=TCGTAGCGGG"

ORIGIN
Alignment Scores:
Pred. No.:      2.34e-75      Length:      720
Score:          875.00      Matches:     171
Percent Similarity: 81.93%      Conservative: 24
Best Local Similarity: 71.85%      Mismatches:  32
Query Match:    28.05%      Indels:     12
DB:              13          Gaps:        2

US-09-980-364-2 (1-579) x B0864461 (1-720)

Qy 222 ArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLysGly 241
Db 2 AGATACGAGGCTCATTTATGGACAAATAGTTGCAGAGAGAGAGCAAGTTCGCAGAGGA 61
Qy 242 ArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAspLeu 261
Db 62 AGACAAGTTTATTTGGGTGGTTACGACAAAGAGATAGGCAGCAGAGCTTATGATTTA 121
Qy 262 AlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyrGlu 281
Db 122 GCTGATTGAATATTTGGGTGGTACCACTACGACAAATTTCCCTATTAGCACTATGAG 181
Qy 282 LysGluValGluGluMetLysHisMetThrArgGlnGlnTyrValAlaSerLeuArgArg 301
Db 182 AAGGAGATCGGAGATGAACACATGACTAGACAAAGATACGTAGCATCTCTTAGAAGG 241

```

```

Qy 302 LysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGln 321
Db 242 AAAAGTAGCGGTTTTTCTCGAGGTGCTTCCATCTATAGAGGTGTGACAGACACCATCAA 301
Qy 322 HisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGly 341
Db 302 CATGGAAGATGGCAAGCAAGAAATGGAAAGAGTTGCAGGGAACAAGATCTTTACTTGGGA 361
Qy 342 ThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAspIleAlaIleLysPheArg 361
Db 362 ACTTTTGGTACACAGGAAGAGCTGCTGAGGCTTATGACATTGCTGCATCAATTCGCT 421
Qy 362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGlu 381
Db 422 GGACTGATGTCAGTCACAAATTTGAAATCAACAGATACGACGTTAAATGCATACTTGAA 481
Qy 382 SerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaA-SenArgProVa 401
Db 482 AGCACCACTTACCGGTTGGTGGTGAACGCTAAACGCTCAAGGATGAGAACAGCCGCT 541
Qy 401 lpro-----SerMetMetMetIleSerAsnAsnValSerGluSe 414
Db 542 GCCACTGATGTTACTTCGGGCTGCAAAAGATGATCAC-TTGACCACCACTGACCCACGA 600
Qy 414 rGluAsnSerAlaSerGly-----TrpGlnAsnAlaAlaValGlnHisHisGlnGln 431
Db 601 GGGACTCAACAGTTATGTCACCATGACATGCGCCGAGCTCCATTAACTACCAACCAAGC 660
Qy 431 yValAspLeuSerLeuLeuHisGlnHisGlnGluArgTyrAsnGlyTyrTyr 448
Db 661 CTTGCTCAAGCACAGGCGCAACCGTTTAGCATGCATATCCACACACTAT 712

RESULT 8
BU964897
LOCUS
DEFINITION
  sac04a07.yl Cm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID:
  Cm-cl036-12685 5' similar to TR:Q9SYC2 Q9SYC2 F11M15.6 PROTEIN. ;
  mRNA sequence.
ACCESSION
  BU964897
VERSION
  BU964897.1 GI:24205644
KEYWORDS
  Glycine max (soybean)
SOURCE
  Glycine max
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
    Glycine.
REFERENCE
  1 (bases 1 to 552)
    Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,
    Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
    Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
    Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
    Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
    McCann, R., Waterston, R. and Wilson, R.
    Public Soybean EST Project
    Unpublished (1999)
    Contact: Shoemaker R/Public Soybean EST Project
    Public Soybean EST Project
    Washington University School of Medicine
    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
    Tel: 314 286 1800
    Fax: 314 286 1810
    Email: est@watson.wustl.edu
    This clone is available through: ResGen, Invitrogen Corp. 2130
    South Memorial Parkway Huntsville, AL 35801 For further information
    call: (800)-533-4363 or contact: ccu@resgen.com web site:
    www.resgen.com
    Seq primer: -40RP from Gibco
    High quality sequence stop: 445.
FEATURES
  source
    1..552
      /organism="Glycine max"
      /mol_type="mRNA"

```

/db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl036-12685"
 /tissue_type="somatic embryos cultured on MSD 20"
 /lab_host="DH10B"
 /clone_lib="Gm-cl036"
 /note="Vector: pSPOR1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPOR1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Alignment Scores:
 Pred. No.: 2,686-75 Length: 552
 Score: 872.50 Matches: 162
 Percent Similarity: 96.72% Conservative: 15
 Best Local Similarity: 88.52% Mismatches: 5
 Query Match: 27.97% Indels: 1
 DB: 13 Gaps: 1

US-09-980-364-2 (1-579) x BU964897 (1-552)

QY 213 GlyValThrArgHisArgTyrGlyValArgTyrGluAlaHisLeuTyrAspAsnSerCys 232
 DB 3 GGAGTAACGGACATAGATGACTGGAGGTATGAGCTCACCTTTGGGATATAGCTGT 62
 QY 233 LysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGlu 252
 DB 63 AGAAGGGAAGGCAATCAAGAAAGGACCCCAAGTTTATTGGGTGATATGATAAGAA 122
 QY 253 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyThrThrThr 272
 DB 123 GAATAAGCAGCTAGTCTTATGCTCCTCACTGAAGTACTGGGGGACATCCACCACT 182
 QY 273 ThrAsnProMetSerGluTyrGluLysGluValGluGluMetLysHisMetThrArg 292
 DB 183 ACCAACTTTCCAAATAGTAATATGATGAGAGGAAATGGATGAATGAACACATGACGCGA 242
 QY 293 GlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIle 312
 DB 243 CAAGAATTGTGTGTCATTAGAAAGAAAGCAAGTGTCTTCCAGGGGTGCATCAATG 302
 QY 313 TyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyVal 332
 DB 303 TATCGTGGAGTTACAGGCATCACACACGGAAGATGGCAAGCAAGAAATTCACAGATT 362
 QY 333 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaGluAla 352
 DB 363 GCAGGAACAAGATCTTTACTTGGGAACCTTCAGTACTGAAGAGAGGGTGTGTAAGCA 422
 QY 353 TyrAspIleAlaAlaLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsn 372
 DB 423 TACGACATAGCTGCATTAAGATTTCAGAGTCTCAACGCTGTCAAACTTTGATGACG 482
 QY 373 ArgTyrAsnValLysAlaLysLeuGluSerProSerLeuProIle---GlySerAlaAla 391
 DB 483 CGCTTACAGCTGAAAGCACTCTTGAAGCAACACTCTCCCAATAGGAGGAGGCTGCA 542
 QY 392 LysArgLeu 394
 DB 543 AAGGCTCTG 551

RESULT 9

BM307997
 LOCUS
 DEFINITION
 sak40e09.y1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl036-5586 5' similar to TR:Q95YC2 Q95YC2 F11M15.6 PROTEIN. ; mRNA sequence.
 BM307997
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE
 1 (bases 1 to 593)
 Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,F., Gibbons,M., Fape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 427.
 Location/Qualifiers
 1..593
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl036-5586"
 /tissue_type="somatic embryos cultured on MSD 20"
 /lab_host="DH10B"
 /clone_lib="Gm-cl036"
 /note="Vector: pSPOR1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPOR1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Alignment Scores:
 Pred. No.: 5,946-75 Length: 593
 Score: 869.50 Matches: 157
 Percent Similarity: 90.10% Conservative: 25
 Best Local Similarity: 77.72% Mismatches: 15
 Query Match: 27.88% Indels: 5
 DB: 12 Gaps: 1

US-09-980-364-2 (1-579) x BM307997 (1-593)

QY 162 LeuSerLeuSerMetAsnSerThrSerCysAspAsnAsnAspSerAspAsnAsn 181


```

654 AGGAGCTGGAGAGATGAAGTCGATGACGGCGGAGGAGTTTCATCGCTCGCTCGTACG 713
302 LysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGln 321
714 AAGAGCAGCGGCTTCTCGGAGGGCATCCATCTACAGAGGAGTAACAGGCGATCATCAG 773
322 HisGlyArgTropGlnAlaGileGlyArgValAlaGlyAsnLysAspLeuTyrLeuGly 341
774 CACGGCGGTGGCAGCAGAGATCGCAGGGTGGCCGGAAACAG-GACCTGTACCTGGGA 832
342 ThrPheGlyThrGlnGluAlaGluAlaGluAlaTyrAspIleAlaAlaIleLysPheArg 361
833 ACTTTTCAGCAGCAGCAGGAGGGCGGAGGGGTACGACATTCGCGCGCATCAAGTTCGCG 892
362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGlu 381
893 GGCTCACGCG-GTGACCAACTC-GACATGAG-CGCTACGAGCTCGAGGATCCTCACA 949
382 SerProSerLeuProIleGlySerAlaAlaLys 392
950 GCGACCTGCCATCGCGCGCGGGCGGCGCGCGCGG 982

RESULT 12
BJ188928
LOCUS
DEFINITION
  BJ188928 normalized full length cDNA library, chloronemata,
  caulonemata and malformed buds Physcomitrella patens subsp. patens
  cDNA clone pphb45d02 5', mRNA sequence.
ACCESSION
  BJ188928
VERSION
  BJ188928.1 GI:18356869
KEYWORDS
  EST.
SOURCE
  Physcomitrella patens subsp. patens
  Physcomitrella patens subsp. patens
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
  Bryopsida; Funariaceae; Funariales; Funariaceae; Physcomitrella.
  1 (bases 1 to 555).
  Nishiyama,T., Fujita,T., Shin-i,T., Seki,M., Nishide,H.,
  Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,
  Kohara,Y. and Hasebe,M.
  Comparative genomics of Physcomitrella patens gametophytic
  transcriptome and Arabidopsis thaliana: implication for land plant
  evolution
  Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
  22709184
  12806149
  Contact: Tadasu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp
  A backbone of the vector is basically from pBluescript II (KS),
  that was in vivo excised from a l-FLC phage vector (Carninci et al.
  2001). 5' end of the cDNA that was digested with XhoI was ligated
  to SalI site of the vector and the 3' end including polyA tail was
  ligated to BamHI site of the
  vector[5].- gagAGAGAGAGATCAACCTGGgAgAgTTTTTTTTTTTTTIVN-3' was
  used as a 1st 3' primer, and
  5'-ggTTCGAGTCATCGTGTCCAGACAGCGATGACTCGAGAACCGNNNN-3' as 2nd
  5'-hairpin primer, giving the following 5' boarder sequence,
  AGCCAAATCGCGAGCTCGAATTCGTCGAGAACCG). cDNA instert could be
  amplified with conventional T7 and T3 primers. This full-length
  cDNA library was generated according to the method described in
  Nishiyama et al. (2003).
  Protonemata were blended by the POLYTRON, and then cultivated on
  the BCDATG medium for 13- 14 days under the continuous light.
  These clones are available from RIKEN Bio Resource Center
  (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database
  of Physcomitrella EST clones is available at the PHYSCOBASE
  (http://moss.nibb.ac.jp);
  Location/Qualifiers

```

```

1. 555
/organisms="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clones="pphb45d02"
/tissue_type="mixture of chloronemata, caulonemata and
malformed buds"
/clone_lib="normalized full length cDNA library,
chloronemata, caulonemata and malformed buds"

source
Alignment Scores:
Pred. No.: 3.4e-72 Length: 555
Score: 841.00 Matches: 150
Percent Similarity: 94.57% Conservative: 24
Best Local Similarity: 81.52% Mismatches: 10
Query Match: 26.96% Indels: 0
Dbs: 12 Gaps: 0

US-09-980-364-2 (1-579) x BJ188928 (1-555)
QY 191 AspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIle 210
Db 4 GACAACAAGATCCGTCCTCCCGCCCAAGTCCATCGATACCTTCGGCCAGAGAATCTGTG 63
QY 211 TyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTyrAspAsn 230
Db 64 TACAGAGCGGTCACTAGGCATCGTTGGACTGGAGCGGTACGAGGCACATCTCTGGACAA 123
QY 231 SerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAsp 250
Db 124 ACTGTAGAAAGGAAGGTCAACTCGCAAGGTGACAAAGTATATCTTAGGAGGATATGAT 183
QY 251 LysGluGluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyrTyrGlyThrThr 270
Db 184 AAAGAAGCAAGGCGAGCCAGAGCTTATGCTGGCAGCGCTCAAAATATTTGGGTTCACG 243
QY 271 ThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLysHisMet 290
Db 244 ACCACCATCACTTCACCTTGATACATACGACAAAGAGCTGGAGAGATGAAGAACA 303
QY 291 ThrArgGlnGluTyrValAlaSerLeuArgLysSerSerGlyPheSerArgGlyAla 310
Db 304 TCTCGCAGGAGTATGTTCCTCGCTTGAGAAAGAAAGCAGTGGCTTTTCGAGAGAGCA 363
QY 311 SerIleTyrArgGlyValThrArgHisGlnHisGlyArgTyrGlnAlaArgIleGly 330
Db 364 TCCATGTATCGAGGGGTGACGAGACCCATCAGCATGGAGATGCGAAGCAGCATTCGG 423
QY 331 ArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluAlaAla 350
Db 424 CGAGTTGCGGGAAACAAGGAGCTGTACCTGGGCACTTACAGCACTCAAGAGGAGCAGCG 483
QY 351 GluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAsp 370
Db 484 GAGGCTACGACATAGCTCCATAAAATATCGTGGAAATATATCCCGTCAAACTTTCAT 543
QY 371 MetAsnArgTyr 374
Db 544 ATCTCTCGTTAC 555

RESULT 13
BJ178045
LOCUS
DEFINITION
  BJ178045 normalized full length cDNA library, chloronemata,
  caulonemata and malformed buds Physcomitrella patens subsp. patens
  cDNA clone pphb21g12 5', mRNA sequence.
ACCESSION
  BJ178045
VERSION
  BJ178045.1 GI:18346002
KEYWORDS
  EST.
SOURCE
  Physcomitrella patens subsp. patens
  Physcomitrella patens subsp. patens
  ORGANISM

```


40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucrest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:
Pred. No.: 9,52e-72 Length: 690
Score: 838.00 Matches: 169
Percent Similarity: 81.36% Conservativeness: 23
Best Local Similarity: 71.61% Mismatches: 26
Query Match: 26.87% Indels: 19
DB: 13 Gaps: 5

US-09-980-364-2 (1-579) x CA094556 (1-690)

QY 129 GlyGlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAsp 148
Db 21 GGGCGAAGCTCTATCGGGCTGTCATGATCAAGAACTGGCTGGCAGTCAGCGCGCCT 80
QY 149 AsnValAspAsnGlnGluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSer 168
Db 81 -----GCACCTCGGGGGTGCAC-----TCCATGGCGCTG 110
QY 169 SerThrSerCysAspAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThr 188
Db 111 GCGACACGCGC---ACATCGCTGAGAGAGCAGGAGGTGTGCACCGCGGGAGAGC 167
QY 189 -----lleAspAspSerValGluAlaThrProLysLysThrIleGluSer 203
Db 168 GTTGGCATCATCTTGTGGACATCGGCAGCAGAGG-----AAGCCCTGGTGGACACA 221
QY 204 PheGlyGlnArgThrSerIleLysArgGlyValThrArgHisArgTrpThrGlyArgTyr 223
Db 222 TTGGGCGAGCGACTTCTATTATTCGGGGGTGCACAAAGCATAGATGGACAGAGGTAT 281
QY 224 GluAlaHisLeuTrpAspAsnSerCysLysArgGlyGlnThrArgLysGlyArgGln 243
Db 282 GAGCCCATCTTTGGGCAATAGCTGCGAAGAGAGTGTGAGTGTGAGTGTGAGTGTGCT 341
QY 244 ValTrpLeuGlyGlyTrpAspLysGluLysAlaAlaArgLysAspLeuAlaAla 263
Db 342 GTATACCTAGTGTGATATGATATAAGAGAGAGAGAGTGTGAGTGTGAGTGTGCT 401
QY 264 LeuLysTrpTrpGlyThrThrThrThrAspPheProMetSerGluTrpGluLysGlu 283
Db 402 CTCAAGTATTGGGCGACTACACCACTACAAATTTTCCATGAGCAACTATGAAGAAGAA 461
QY 284 ValGluGluMetLysHisMetThrArgGlnGluTrpValAlaSerLeuArgArgLysSer 303
Db 462 CTGGAAGAGATGAGCATATGTTCAGCAAGAAATATGTTCGCTGTAGAGGAAAGC 521
QY 304 SerGlyPheSerArgGlyAlaSerIleLysArgGlyValThrArgHisGlnHisGly 323
Db 522 AGTGGATTTCTCGTGTGTCATCGATTACCGAGGGTGTACCGAGCCACCGATGGA 581
QY 324 ArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTrpLeuGlyThrPhe 343
Db 582 GG-TGGCAACGAGATAGAGAGAGTGGCAGAAACAGGATCTGTTATTTAGGCACCATC 640
QY 344 GlyThrGlnGluAlaAlaGluAlaTrpAspIleAlaAlaLys 359
Db 641 AGTACCAGGAAGAGCGGAGAGCTTATGACATCGTGTGCTGCAACAA 698

RESULT 15
AW780688 500 bp mRNA linear EST 03-DEC-2001
LOCUS sl75e07.y1 Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl027-7165 5' similar to TR:Q41832 Q41832 OPEN READING FRAME. ?;
mRNA sequence.
ACCESSION AW780688
VERSION AW780688.1 GI:7795291
KEYWORDS EST.

SOURCE ORGANISM

Glycine max (soybean)
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE AUTHORS

1 (bases 1 to 500)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.

TITLE JOURNAL COMMENT

Public Soybean EST Project
Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 412.

FEATURES source

Location/Qualifiers
1..500
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-7165"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/clone_lib="Gm-cl027"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First-strand synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer (GAGAGAGAGAGAGAGAGAGAGTGTCTCGAG(T)18) to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in
with cloned Pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The XhoI site within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500 bp cutoff,
using GibcoBRL Life Technologies' cDNA Size Fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript(tm) II XR Predigested vector
(pBluescript II SK+) that has been digested with EcoRI
and XhoI, and phosphorylated by Stratagene). 97% of the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=30). This
library was constructed by Dr. Paul Keim and Dr. Virginia
Coryell."

ORIGIN

Alignment Scores:
Pred. No.: 1.4e-71 Length: 500
Score: 834.00 Matches: 154
Percent Similarity: 97.59% Conservativeness: 8

Best Local Similarity: 92.77% Mismatches: 4
Query Match: 26.74% Indels: 0
DB: 10 Gaps: 0

US-09-980-364-2 (1-579) x AW780688 (1-500)

```
QY 206 GlnArgThrSerIleTyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAla 225
Db 3 GAGCGCACCTCCATCTTACCGCGCGCTACCCGACATAGATGCGGGAAGATACGAAGCT 62
QY 226 HisLeuTyrAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyr 245
Db 63 CATCTATGGACAATAGATTGTAGAGAGAAAGCCAAAGCAGGAAAGGAAAGACAAGTTTAC 122
QY 246 LeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLys 265
Db 123 CTGGGTGTTATGACAAGAAAGTAAGGAGCCAGCGGCTTACGATCTCGCAGCTCTCAAG 182
QY 266 TyrTrpGlyThrThrThrAsnPheProMetSerGluTyrGluLysGluValGlu 285
Db 183 TACTGGGTCCCAACTACCACCACCACTTCCCAATTCCAACTATGACAGGAACTGGAG 242
QY 286 GluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGly 305
Db 243 GAGATGAAGAACATGACCAGGCAAGATTGTGCTTCTACGAAAGGAGAGAGTGGT 302
QY 306 PheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGlyArgTyr 325
Db 303 TTCTCTAGGGGGCCCTCTATATACAGAGAGTGAGGAGACACACCAGCATGGCCGATGG 362
QY 326 GlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThr 345
Db 363 CAGCGAGAATAGGAGAGTTGCCGGAACAAAGACCTCTACCTTGGAACTTTCAGCACC 422
QY 346 GlnGluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAla 365
Db 423 CAAGAAAGAGCTGCTGAGCGCTATGACATTCGCTATCAAAATTCAGGGGATTAATGCA 482
QY 366 ValThrAsnPheAspMet 371
Db 483 GTCACAAACTTTGACATG 500
```

Search completed: July 19, 2004, 18:27:44
Job time : 4169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 10:22:57 ; Search time 6533 Seconds
(without alignments)
3841.362 Million cell updates/sec

Title: US-09-980-364-2
Perfect score: 3119
Sequence: 1 MNNWLGFLSPYEQNHRRK.....YHGGGGEVAPFTVWVDN 579

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09980364/runat_15072004_095235_25866/app_query.fasta_1.775
-DB=GenEmbl -QFMT=fastap -SUFFIX=ige -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=DCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09980364@cgn2_1_4198 @runat_15072004_095235_25866 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hug.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3119	100.0	2014	6	AX058687 Sequence
2	3119	100.0	2014	8	AF317904 Brassica
3	3041	97.5	1992	8	AF317905 Brassica
4	3041	97.5	2011	6	AX058689 Sequence
5	2690.5	86.3	4873	6	AX058691 Sequence
6	2690.5	86.3	4873	6	AX058691 Sequence
7	2096	67.2	33563	8	AF317906 Brassica
8	2073.5	66.5	5151	6	AX058692 Sequence
9	2073.5	66.5	5151	8	AF317907 Arabidops
10	1008	32.3	2546	8	AY461432 Nicotiana
11	994	31.9	2153	8	AK111891 Oryza sat
12	973.5	31.2	1926	6	AX555220 Sequence
13	973.5	31.2	2510	8	AK106306 Oryza sat
14	965.5	31.0	2164	8	AY062179 Oryza sat
15	963	30.9	1745	8	AK101959 Oryza sat
16	942	30.2	2010	6	AX555223 Sequence
17	942	30.2	2344	6	AX555216 Sequence
18	930	29.8	1773	8	ZMWHC1
19	924	29.8	2323	6	AX555218 Sequence
20	913.5	29.3	1738	6	BD274518 Methods f
21	913.5	29.3	1738	6	AR316369 Sequence
22	913.5	29.3	1738	6	AR427903 Sequence
23	891	28.6	1699	8	AY117207 Arabidops
24	891	28.6	1905	8	ATU41339 Arabidops
25	891	28.6	2056	8	AY080706 Arabidops
26	891	28.6	2148	6	BD274516 Methods f
27	891	28.6	2148	6	AR316367 Sequence
28	891	28.6	2148	6	AR427901 Sequence
29	891	28.6	2148	8	ATU40256 Arabidops
30	891	28.5	1955	8	ATU44028 Arabidops
31	809	25.9	1660	8	AK109839 Oryza sat
32	798	25.6	105807	8	AC006085 Arabidops
33	771	24.7	1749	6	AX048247 Sequence
34	752	24.1	81677	8	AB025629 Arabidops
35	744	23.9	153439	8	AP003313 Oryza sat
36	741.5	23.8	94788	2	AC144930 Medicago
37	727.5	23.3	1440	8	AK109848 Oryza sat
38	709.5	22.7	4024	8	AY062180 Oryza sat
39	685	22.0	1069	8	AY133880 Arabidops
40	685	22.0	1237	8	AY045915 Arabidops
41	677	21.7	103788	8	AP004747 Oryza sat
42	674.5	21.6	1572	8	AK111617 Oryza sat
43	670.5	21.5	110766	8	ATT28119 Arabidops
44	670.5	21.5	198493	8	ATCHRIV88 Arabidops
45	667	21.4	1866	8	AK106769 Oryza sat

ALIGNMENTS

RESULT 1

AX058687	AX058687	2014 bp	DNA	linear	PAT 17-JAN-2001
LOCUS	Sequence 1 from Patent WO0075330.				
DEFINITION	AX058687				
ACCESSION	AX058687.1	GI:12311028			
VERSION					
KEYWORDS	Brassica napus (rape)				
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	Boutiller, K., Ouellet, T., Custers, J., Hattori, J., Miki, B. and van Loocker Campagne, M.				
TITLE	Use of the bnm3 transcriptional activator to control plant embryogenesis and regeneration processes				
JOURNAL	Patent: WO 0075330-A1 14-DEC-2000; Plant Research International (NL); THE MINISTER OF AGRICULTURE (CA)				
FEATURES	Location/Qualifiers				
source	1..2014				
ORIGIN	/organism="Brassica napus" /mol_type="unassigned DNA" /db_xref="taxon:3703"				
Alignment Scores:					
Pred. No.:	1..61e-214	Length:	2014		
Score:	3119.00	Matches:	579		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-980-364-2 (1-579) x AX058687 (1-2014)					
QY	1 MetAsnAsnTrrLeuGlyPheSerLeuSerProTrrGluGlnAsnHisHisArgLys	20			
DB	120 ATGAATAATACTGGTGGAGGCTTCTCTCTCTCTATGAAACAAATCACCATCGTAAG	179			
QY	21 AspValTrrSerSerThrThrThrValValAspValAlaGlyGluTrrCysTrrAsp	40			
DB	180 GACGCTACTCTTCACCAACACACGCTGTAGATGTCGCGGAGAGTACTGTACGAT	239			
QY	41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly	60			
DB	240 CCGACCGCTGCTCGATGAGTCTTCAGCCATCCAAACATCGTTCTCTCTCTCTGTT	299			
QY	61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrrPaspilleAsn	80			
DB	300 GTCGTCGTCGATGCTTTCACCAAGAGACACAAATAGTCTACTCCCGAGATTGGGACATCAAT	359			
QY	81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPhe	100			
DB	360 GGTGTGTCATCAATAACATCCCAACGATGAGCAGATGACCAAGCTTGAGAAATTC	419			
QY	101 LeuGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr	120			
DB	420 CTTGGCCGACACACGATTTACACACCAACCAACGATTTGGAGATGGAGTGGAGT	479			
QY	121 GlyCysTrrGlyGlyAspGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly	140			
DB	480 GGCTGTTATGAGGAGGAGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	539			
QY	141 TrrLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys	160			
DB	540 TGGCTGAGAAATCAACCCGCGGATTAATGTTGATATCAAGAAATGCAATGCTGCAAAA	599			
QY	161 GlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsnAsn	180			
DB	600 GGCCTGTCTCTCAATGAACTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	659			
QY	181 AsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrProLysLysThr	200			
DB	660 AACGTTGTTGCCCAAGGAGACATATTGATGATAGCGTTGAGCTTACACCGAAGAAACT	719			
QY	201 IleGluSerPheGlyGlnArgThrSerIleTrrArgGlyValThrArgHisArgTrrPthr	220			
DB	720 ATTGAGAGTTTGGACAGAGGAGCTCTATATACCGCGGTGTACAAAGGATCGGTGGACA	779			
QY	221 GlyArgTrrGluAlaHisIleuTrrPaspAsnSerCysLysArgGluGlyGlnThrArgLys	240			
DB	780 GGAAGATATGAGGACATTTATGGGATAATAGTTGTAAAGAGAGAGGCCAAACGCGCAAA	839			
QY	241 GlyArgGlnValTrrLeuGlyGlyTrrAspLysGluGluLysAlaAlaArgAlaTrrAsp	260			
DB	840 GGAAGACAAAGTTTATTTGGAGGTTATGACAAAGAAAGAAAAGCAGCTTAGGCTTATGAT	899			
QY	261 LeuAlaAlaLeuLysTrrTrrPglyThrThrThrThrThrThrThrThrThrThrThr	280			
DB	900 TTAGCCCGCATCAAGTATTGGGAAACCCACCTACTACTTAACCTCCCATGAGCGAATAT	959			
QY	281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTrrValAlaSerLeuArg	300			
DB	960 GAAAGAGGTAGAGAGATGAAGCACATGACAGCAAGCAAGATATGTTGCCCTCACTCGGC	1019			
QY	301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTrrArgGlyValThrArgHisHis	320			
DB	1020 AGGAAAAGTAGTGGTTCT	1079			
QY	321 GlnHisGlyArgTrrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTrrLeu	340			
DB	1080 CACATGGAAGATGGCAAGCTAGGATAGGAGAGTGGCGGTAAACAAGACCTCTACTTG	1139			
QY	341 GlyThrPheGlyThrGlnGluAlaAlaGluAlaTrrAspIleAlaAlaIleLysPhe	360			
DB	1140 GGAACCTTTTGGCACACAAGAAAGCTGCAGAGGATACGACATTCGCGCCATCAAAATC	1199			
QY	361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTrrAsnValLysAlaIleLeu	380			
DB	1200 AGAGATTAACCGCAGTGAATTAACCTTCAGCATGAACAGATCAACAGTTAAAGCAATCTC	1259			
QY	381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro	400			
DB	1260 GAAAGCCTTAGTCT	1319			
QY	401 ValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGly	420			
DB	1320 GTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1379			
QY	421 TrrGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGlnHis	440			
DB	1380 TGGCAAAACGCTGCGGTTTCAGCATCATCAGGAGTAGATTTGAGCTTATTGCACCAAT	1439			
QY	441 GlnGluArgTrrAsnGlyTrrTrrTrrTrrTrrTrrTrrTrrTrrTrrTrrTrrTrr	460			
DB	1440 CAAGAGAGTACAAATGGTATTATTATCAATGGAGGAAACTTGTCTTCGAGAGTCTAGG	1499			
QY	461 AlaCysPheLysGlnGluAspAspGlnHisIlePheLeuSerAsnThrGlnSerLeuMet	480			
DB	1500 GCTTGTTCACCAAGAGAGATGATCAACCATTTCTTTCAGCAACACGACGAGCTCATG	1559			
QY	481 ThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnVal	500			
DB	1560 ACTAATATCGATCATCAAAAGTTCTGTTTCGGATGATTCGGTTACTGTTGTGGAAATGT	1619			
QY	501 ValGlyTrrGlyGlyTrrGlnGlyPheAlaAlaProValAsnCysAspAlaTrrAlaAla	520			
DB	1620 GTTGTATTGTTGTTTATCAAGGATTTTCAGCCCGGTTAACTCGATGCTTCGCTGCT	1679			
QY	521 SerGluPheAspTrrAsnAlaArgAsnHisTrrTrrPheAlaGlnGlnGlnThrGln	540			
DB	1680 AGTGATTTGATTAACGCAAGAAACCATTTACTTTTCTCAGCAGCAGCAGCAGCAG	1739			
QY	541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTrr	560			

Db 1740 CAGTCCAGGTGGAGATTTTCGGCGCAATGACGAATAATGTTGGCTCTAATATGTAT 1799

Qy 561 TyrHisGlyGluGlyGluValAlaProThrPheThrValTyrAsnAspAsn 579

Db 1800 TACCATGGGAAGGTGGTGGAGAGTTGCTCCACATTTACAGTTTGGACACCAAT 1856

RESULT 2

AF317904 2014 bp mRNA linear PLN 10-OCT-2002

LOCUS Brassica napus AP2/EREBP transcription factor BABY BOOM1 (BBM1)

DEFINITION mRNA, complete cds.

ACCESSION AF317904

VERSION AF317904.1

KEYWORDS GI:21069050

SOURCE Brassica napus (rape)

ORGANISM Brassica napus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 2014)

AUTHORS Boutillier, K., Offringa, R., Sharma, V.K., Kieft, H., van Lammeren, A.A.M., Ouellet, T., Zhang, L., Hattori, J., Liu, C.-M., Miki, B.L.A., Custers, J.B.M. and van Lookeren Campagne, M.M.

TITLE Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth

JOURNAL Plant Cell 14 (8), 1737-1749 (2002)

MEDLINE 22163009

PUBMED 12172019

REFERENCE 2 (bases 1 to 2014)

AUTHORS Boutillier, K., Ouellet, T. and Hattori, J.

TITLE Direct Submission

JOURNAL Submitted (02-NOV-2000) Business Unit Plant Development and Reproduction, Plant Research International, Droeendaalsesteeg 1, Wageningen 6708 PB, The Netherlands

FEATURES

source

1..2014

/organism="Brassica napus"

/mol_type="mRNA"

/cultivar="Topas"

/db_xref="taxon:3708"

1..2014

/gene="BBM1"

120..1859

/gene="BBM1"

/codon_start=1

/product="AP2/EREBP transcription factor BABY BOOM1"

/protein_id="AA03800.1"

/db_xref="GI:21069051"

/translation="MNNWLGFSLSPEYQNHHRKDYSSTTTTTVDVAGEYCYDPTAA

SDSSATQTFPPFVGVVDAFTDRDNNSHRWDINGCCACNNIHNDQCPKLENFLG

RTTIYNTNENVGSGGCGGGLSLSMIKTLRNOPVDNVDNQENNAK

GLSLMNSSTSCDNNDSNNVVAQGTIDDSVEATPKTIESFGORTSIYRGVTRHP

WTRYEAHLNDNCKREGQTRKGRQVYLGVDKEEKAAYDLAALKYWGTTTNNFP

MSYEKEVEEMKMTROYEVASLRKSSGFSRGSASYRGVTRHHQHRQALGRVAG

NKDLYLGTFTQEEAAYDIAAKFRGLTAVTNFDNNVYNVKAILESPLPISAAK

RLKEANRPVPMIMNNVSESNASQWQNAVOHQGVDSLHLHQHRYNGYYN

GLSSSASACFKQDDQHFLSNLQSLMTNIDHQSVDSDSVTVGVGVGYGQ

FAFPVNDVAAAEFDYARNHYVFAQQQQTQSPGDFPAAATNNVGSNNVYHGGEG

GEVAPTIVNDN"

ALIGNMENT SCORES

Pred. No.: 1,61e-214 Length: 2014

Score: 3119.00 Matches: 579

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 8 Gaps: 0

US-09-980-364-2 (1-579) x AF317904 (1-2014)

Qy 1 MetAsnAsnAnTrrPLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLys 20

Db 120 ATGAATAATAACTGGTTAGGCTTTTCTCTCTCTCTATGACAAAATCACCATCGTAAG 179

Qy 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40

Db 180 GACGCTACTCTCCACCACCACCAACCGTGGTAGATGTCGCGGAGAGTACTGTTAGCAT 239

Qy 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60

Db 240 CCGACCGGTGCTCCGATGAGTCTTCAGCCATCCAAACATCGTTCTCTCTCTCTTGGT 299

Qy 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTTPAspIleAsn 80

Db 300 GTCGTCGATGCTTTCACACAGAGACAACATAGTCACTCCCGAGATTGGGACATCAAT 359

Qy 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPhe 100

Db 360 GGTGTGTCATCAATAACATCCACACGATGAGCAAGATGACCAAGCTTGAAATTC 419

Qy 101 LeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySer 120

Db 420 CTTGGCCGACACACACGATTTACACCAACCAACCAACGTTGGAGATGGAAGTGAAGT 479

Qy 121 GlyCysTyrGlyGlyAspGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 140

Db 480 GGTGTTATGAGAGAGAGAGCGGTGGTGGTCACTAGGACTTTTCGATGATAAGACA 539

Qy 141 TrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys 160

Db 540 TGGCTGAGAAATCAACCCGCGATGATGTTGATATCAAGAAATGGCAATGCTGCAAAA 599

Qy 161 GlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsn 180

Db 600 GGCCTGTCTCTCAATCACTCATCTACTTCTTGTGATAACAACAACAGACGCAATAAC 659

Qy 181 AsnValValAlaGlnGlyThrIleAspAspSerValGluAlaThrProLysThr 200

Db 660 AACGTTGTGCCCAAGGAGAGACTATTGATGATGCGTTGAAGCTACACCGAGAAACT 719

Qy 201 IleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTyr 220

Db 720 ATTGAGATTTGGACAGAGAGCGTCTATATACCGCGTGTACAGGCGATCGTGGACA 779

Qy 221 GlyArgTyrGluAlaHisLeuTyrAspAsnSerCysLysArgGluGlyGlnThrArgLys 240

Db 780 GGAAGATATGAGCACATTTATGGGATATAGTTGTAAGAGAGAGGCGCAACCGCGCAA 839

Qy 241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAsp 260

Db 840 GGAAGACAGATTTATTTGGAGGTTATGCAAAAGAGAAAGAGAGCTAGGCTTATCAT 899

Qy 261 LeuAlaAlaLeuLysTyrTyrGlyThrThrThrThrAsnPheProMetSerGluTyr 280

Db 900 TTAGCGGCACCTCAAGTATTGGGAACCACTACTACTACTACTACTACTACTACTACT 959

Qy 281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300

Db 960 GAAAAGAGGTGAGAGAGATGAGACATGACAGGCAAGAGTATGTTGCTCTCATCGCC 1019

Qy 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320

Db 1020 AGGAAAAGTAGTGGTTTCTCTCGTGGTCAATGATTTATCGTGGAGTAAACAGACATCAC 1079

Qy 321 GlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 340

Db 1080 CAACATGGAAGATGCGAGCTAGGATAGAGAGTCCCGCGGTAAACAAGACCTCTACTTG 1139

Qy 341 GlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaLysPhe 360

Db 1140 GGAACCTTTGGCACACAGAGAGAGCTGCAGAGGATACGACATTCGCGGCATCAATTC 1199

Qy 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaLeu 380

Db 1200 AGAGGATTAAACCGCATGTACTAACTTCGACATGAAACAGATACACAGCTTAAACAGTTC 1259

Db 651 ATTGAGATTTGGACAGAGGAGCTCTATATACCGGGTGTACAGGCATCGGTGGACA 710
QY 221 GlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGlnThrArgLys 240
Db 711 CGAAGATATGAGCACATTTATGGGATAATAGTTGTAAACGAGAGGCGCAACGCGCAA 770
QY 241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAsp 260
Db 771 GGAAGCAAGTTTATTTGGAGGTTATGCAAGAGAAAAGCAGCTAGGCGCTTATGAT 830
QY 261 LeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyr 280
Db 831 TTAGCGCACTCAAGTATTTGGGAACACCACTACTACTTAACTTCCCATGAGCGAATAT 890
QY 281 GluLysGluValGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300
Db 891 GAGAAAGAGATAGAAGAGATGAAGCAGCATGCAAGGCAAGATGTTTCCCTCACTTCGC 950
QY 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHis 320
Db 951 AGCAAAAGTAGTGTCTCTCGTGGTGCATCGATTTATCGTGGAGTAAACAAGACATCAC 1010
QY 321 GlnHisGlyArgTrpGlnAlaAlaGlyArgValAlaGlyAsnLysAspLeuTyrLeu 340
Db 1011 CAACATCGAAGATGGAAGCTAGGATAGGAAGAGTCCCGGTAAACAAGACCTCTACTTG 1070
QY 341 GlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPhe 360
Db 1071 GGAACCTTTGGCACACAGNAGAGCTGCAGAGCATACGACATTCGGGCCATCAATTC 1130
QY 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeu 380
Db 1131 AGAGGATTAAACCCAGTGAATACTTCGACATGACAGATACAACTGTTAAAGCAATCTTC 1190
QY 381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro 400
Db 1191 GAAAGCCCTAGTCTTCTATTTAGTGGCCGCCAAACGCTCCAGAGAGCTAACCCGTCG 1250
QY 401 ValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGly 420
Db 1251 GTTCCAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1310
QY 421 TrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeuHisGlnHis 440
Db 1311 TGGCAAAACGCTCGGTTACAGCATCATCAGGAGTAGATTGAGCTTATTTGCGAGCAACAT 1370
QY 441 GlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArg 460
Db 1371 CAAGAGAGGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1430
QY 461 AlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMet 480
Db 1431 GCATTGTTTCAACAAGAGGATGATCAACACCATTTCTTGTGACACACGACGAGCCCTCATG 1490
QY 481 ThrAsnIleAspHisGlnSerSerValSerAspAspSerValThrValCysGlyAsnVal 500
Db 1491 ACTAATATCGATCATCAAGATTCGTTTTCAGATGATTCGTTTTCGTTTTCGTTTTCGTTT 1550
QY 501 ValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAlaAla 520
Db 1551 GTTGTTATGTTGTTATCAAGGATTTGACGCCCGGTTAACTGCGATGCTACGCTGCT 1610
QY 521 SerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
Db 1611 AGTAGTTTGTACTATAACGCAAGAAACCATTTATTTACTTTTGTCTAGCAGCAGACCCAG 1670
QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
Db 1671 CATTCGCCAGGAGGAGATTTTCCCGGCAATGACGAATAATGTTGGCTCTAATATGAT 1730
QY 561 TyrHisGlyGluGlyGlyGluValAlaProThrPheThrValTrpAsnAspAsn 579
Db 1731 TACCATGGGGAAGGTGGTGAAGAAGTTGCTCCAACTTTACAGTTTGAACAGCAAT 1787

RESULT 4
AX058689 AX058689 2011 bp DNA linear PAT 17-JAN-2001
LOCUS Sequence 3 from Patent WO0075330.
DEFINITION AX058689
ACCESSION AX058689
VERSION AX058689.1 GI:12311029
KEYWORDS Brassica napus (rape)
SOURCE Brassica napus
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1
Boutlier,K., Ouellet,T., Custers,J., Hattori,J., Miki,B. and van
lookeren Campagne,M.
Use of the bnm3 transcriptional activator to control plant
embryogenesis and regeneration processes
Patent: WO 0075330-A 3 14-DEC-2000;
Plant Research International (NL) ; THE MINISTER OF AGRICULTURE
(CA)
FEATURES
Location/Qualifiers
source 1..2011
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"
ORIGIN
Alignment Scores:
Pred. No.: 6,32e-209 Length: 2011
Score: 3041.00 Matches: 565
Percent Similarity: 98.45% Conservative: 5
Best Local Similarity: 97.58% Mismatches: 9
Query Match: 97.50% Indels: 0
DB: 6 Gaps: 0
US-09-980-364-2 (1-579) x AX058689 (1-2011)
QY 1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
Db 51 ATGAATAATAAATCTGGTTAGGCTTTTCTCTCTCTTATGAACAAAATCAACATCGTAAG 110
QY 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
Db 111 GACGTCTGCTCTTCCACCACCAACCCCGCTAGATGTCGCGGAGAGTACTGTTACGAT 170
QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
Db 171 CCGACCGCTGCTCGGATGAGTCTTCAGCCATCCAAACATCGTTTCTCTCCCTTTGGT 230
QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsn 80
Db 231 GTCTGCTCTGATGCTTTTCCACGAGACAAACAATAGTCACTCCCGAGATTGGACATCAAT 290
QY 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPhe 100
Db 291 GGTAGTGATGTAGGAGGAGAGCGTGGTGGTCTCACTAGGACTTCGATGATAAAGACA 350
QY 101 LeuGlyArgThrThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySer 120
Db 351 CTTCGCGCGCACCAACCATGATTTTACCAACCAACCAACCAACCAACCAACCAACCAACCA 410
QY 121 GlyCysTyrGlyGlyAspGlyGlyGlyGlySerLeuGlyLeuSerMetIleLysThr 140
Db 411 GGGTGTATTGAGGAGGAGAGCGTGGTGGTCTCACTAGGACTTCGATGATAAAGACA 470
QY 141 TrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys 160
Db 471 TGGCTGAGAAATCAACCCCGTGGATATGTTGATATCAAGAAAATGGCAATCGTGCAAAA 530
QY 161 GlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSerAsnAsn 180
Db 531 GGCTGTCTCTCTCAATGAACTCACTACTTCTTGTGATAACAACAACACTACAGCAGTAC 590

QY 181 AsnValValAlaGlnGlyLysThrIleAspSerValGluAlaThrProLysLysThr 200
 Db : : : : :
 QY 591 AACCTGTTGGCCAAAGGAAGACTATTGATGATACGTTGAAGCTACACCGAAGAACT 650
 Db : : : : :
 QY 201 IleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTrpThr 220
 Db : : : : :
 QY 651 ATTGAGAGTTTGGACAGAGGAGCTATATACCGCGGTGTATACAGGCGATCGGTGGACA 710
 Db : : : : :
 QY 221 GlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLys 240
 Db : : : : :
 QY 711 GGAAGATATAGGCACATTTATGGGATATATAGTTGTAACAGAGGCGCAACCGGCAAA 770
 Db : : : : :
 QY 241 GlyArgGlnValTyrLeuGlyLysTyrAspLysGluGlyLysAlaAlaArgAlaTyrAsp 260
 Db : : : : :
 QY 771 GGAAGACAAAGTTTATTGGGAGGTTATGACAAAGAAAGAAAGCAGCTAGGGCTTATGAT 830
 Db : : : : :
 QY 261 LeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyr 280
 Db : : : : :
 QY 831 TTAGCCGCACCTCAAGTATTGGGAACACCACTACTACTAATCTTCCCATGAGCGAATAT 890
 Db : : : : :
 QY 281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300
 Db : : : : :
 QY 891 GAGAAAGATAGAGAGATGAGACATGACAGGCAAGAGTATGTTGCCTCACTTCGC 950
 Db : : : : :
 QY 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320
 Db : : : : :
 QY 951 AGAAAGTAGTGGTTCTTCGTGGTGCATCGATTATCGTGAGTAACAAGACATCAC 1010
 Db : : : : :
 QY 321 GlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 340
 Db : : : : :
 QY 1011 CAACATGAGATGCAAGTAGGATAGAGAGTCGCGGTACAAAGACCTTACTTTG 1070
 Db : : : : :
 QY 341 GlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaLysPhe 360
 Db : : : : :
 QY 1071 GGAACCTTTTGGCACAAAGAAAGAGCTGCAGAGGCATACGACATTGGCGCCATCAAAATC 1130
 Db : : : : :
 QY 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaLysLeu 380
 Db : : : : :
 QY 1131 AGAGGATTAAACCGAGTGAATCTTCGACATGACATACACGTTAAGCANTCCTC 1190
 Db : : : : :
 QY 381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro 400
 Db : : : : :
 QY 1191 GAAAGCCCTAGTCTCTCTATTGGTAGCGCCGCAAAAGCTCTCAAGGAGGCTAACCGTCG 1250
 Db : : : : :
 QY 401 ValProSerMetMetIleSerAsnValSerGluSerGluAsnSerAlaSerGly 420
 Db : : : : :
 QY 1251 GTTCCAGTATGATGATGATCAGTAATAACGTTTCAGAGAGTGAGATAATGCTAGCGGT 1310
 Db : : : : :
 QY 421 TrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGlnHis 440
 Db : : : : :
 QY 1311 TGGCAAAACCGTCGGTTCACATCATCAGGAGTAGATTGAGCTATTGACGACACAT 1370
 Db : : : : :
 QY 441 GlnGluArgTyrAsnGlyTyrTyrAsnGlyCysAsnLeuSerSerGluSerAlaArg 460
 Db : : : : :
 QY 1371 CAAGAGAGTACAAATGGTTATTATTAATGAGAGAAACTGTCTTCGGAGAGTGCTAGG 1430
 Db : : : : :
 QY 461 AlaCysPheLysGlnGluAspAspGlnHisPheLeuSerAsnThrGlnSerLeuMet 480
 Db : : : : :
 QY 1431 GCTTGTTCAAACAAGAGGATGATCAACACCATTTCTTTGAGCAACACGACGACCTCATG 1490
 Db : : : : :
 QY 481 ThrAsnIleAspHisGlnSerSerValSerAspAspSerValThrValCysGlyAsnVal 500
 Db : : : : :
 QY 1491 ACTAATATCGATCATCAAGATCTCTGTTTCAGATGATTTCGTTTCTGTTTGGAAATGTT 1550
 Db : : : : :
 QY 501 ValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAlaAla 520
 Db : : : : :
 QY 1551 GTTGGTTATGTTGTTATCAAGGATTTCAGCCCGGTTAACTCGGATGCTACGCTGCT 1610
 Db : : : : :
 QY 521 SerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
 Db : : : : :
 QY 1611 AGTGAGTTTGACTATAACGCAAGAAACCATTTACTTTGCTCAGCAGCAGCAGACCCAG 1670

QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
 Db : : : : :
 QY 1671 CATTCGCCAGGAGGAGATTTTCCCGCGGCAATGACGAATATGTTGGCTCTTAATATGAT 1730
 Db : : : : :
 QY 561 TyrHisGlyGluGlyGlyGlyGluValAlaProThrPheThrValTrpAsnAspAsn 579
 Db : : : : :
 QY 1731 TACCATGGGAGAGTGTGGAGAGTGTCTCCAAACATTACAGTTTGGAAACGACAA 1787
 Db : : : : :
 RESULT 5
 AX058691
 LOCUS AX058691 4873 bp DNA linear PAT 17-JAN-2001
 DEFINITION Sequence 5 from Patent WO0075330.
 ACCESSION AX058691
 VERSION AX058691.1 GI:12311030
 KEYWORDS
 SOURCE Brassica napus (rape)
 ORGANISM Brassica napus
 Brassicaceae; Brassicaceae; Brassica.
 REFERENCE 1
 AUTHORS Boutiller,K., Ouellet,T., Custers,J., Hattori,J., Miki,B. and van Lookeren Campagne,M.
 TITLE Use of the 'bmi3 transcriptional activator to control plant embryogenesis and regeneration processes
 JOURNAL Patent: WO 0075330-A 5 14-DEC-2000;
 Plant Research International (NL); THE MINISTER OF AGRICULTURE (CA)
 FEATURES
 Location/Qualifiers
 1..4873
 /organism="Brassica napus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3708"
 misc_feature
 1620..1622
 /note="start codon"
 intron
 1846..2298
 intron
 2720..2952
 intron
 3036..3160
 intron
 3170..3314
 intron
 3404..3553
 intron
 3628..3797
 intron
 3849..3961
 intron
 4039..4148
 intron
 4856..4858
 misc_feature
 /note="stop codon"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,61e-183 Length: 4873
 Score: 2690.50 Matches: 574
 Percent Similarity: 53.24% Conservative: 1
 Best Local Similarity: 53.15% Mismatches: 4
 Query Match: 86.28% Indels: 505
 DB: 6 Gaps: 8
 US-09-980-364-2 (1-579) x AX058691 (1-4873)
 QY 1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLys 20
 Db : : : : :
 QY 1620 ATGAATATACTGTTAGGCTTTTCTCTCTCTTATGAACAAATACCATCTGTAAG 1679
 Db : : : : :
 QY 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40.
 Db : : : : :
 QY 1680 GACGCTACTCTTCCACCACCAACCGCTGATAGTGTCCGCGGAGAGTACTGTACGAT 1739
 Db : : : : :
 QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
 Db : : : : :
 QY 1740 CCGACCGCTGCTCCGATGAGTCTTCAGCCATCAACATCGTTCTCTCTCTCTCTCTCT 1799
 Db : : : : :
 QY 61 ValValValAspAlaPheThrArgAspAsnSerHisSerArg----- 75
 Db : : : : :
 QY 1800 GTCGTCGTCGATGCTTTTCAACAGACAGACAAATAGTCACCTCCCGAGGTATTGTTTGTAGA 1859

Db 4017 GACCTCTACTTGGGAAC-TTTTGGTACGTTTAGTCTCTCTTACTAAACTTCAACATCAA 4075
 QY 343 ----- 343
 Db 4076 ATCTATACAAAGATATCACTAAATAACTACACATATATCTAAGTAAGCTGTACATAT 4135
 QY 344 -----GlyThrGlnGluAlaGluAlaGluAlaGluAlaGluAlaGluAlaGluAla 4135
 Db 4136 ATTATATATGAAGGCACACAGAGAAGCTGCAGAGGCATACGACATTCGCGCCATCAA 4195
 QY 360 PheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValValAlaIle 379
 Db 4196 TTCAGAGGNTTAAACCGAGTGACTTCTGACATGAACATACATACGTTTAAAGCAATC 4255
 QY 380 LeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArg 399
 Db 4256 CTCGAAGACCTTAGTCTCTTCTTGTAGCGCGCAAAACGCTCAAGGAGCTTAACCGT 4315
 QY 400 ProValProSerMetMetIleSerAsnValSerGluSerGluAsnSerAlaSer 419
 Db 4316 CCGGTTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4375
 QY 420 GlyTyrGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGln 439
 Db 4376 GGTTCGCAAAACGCTGCGGTTGAGCATCATCAGGAGTAGATTGAGCTTATTCACCAA 4435
 QY 440 HisGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAla 459
 Db 4436 CATCAAGAGAGTACAAATGCTTATTAATTAATCAATGAGGAACTTGCTCTCGAGAGTCT 4495
 QY 460 ArgAlaCysPheLysGlnGluAspGlnHisPheLeuSerAsnThrGlnSerLeu 479
 Db 4496 AGGGCTGTTTCAACACAGAGAGTATGATCAACACATTTCTTGACACACACGACAGCCTC 4555
 QY 480 MetThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsn 499
 Db 4556 ATGACTAATATCGATCATCAAAAGTCTCTTTCGGATGATTCGGTACTGTTTGTGGAAT 4615
 QY 500 ValValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAla 519
 Db 4616 GTTGTGTTGTTATGTTGTTATCAAGGATTTCAGCCCGGTTAACTCGGATGCTACGCT 4675
 QY 520 AlaSerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThr 539
 Db 4676 GCTAGTGAATTTGATTATACCGAAGAACCATTAATTACTTTGCTCAGCAGCAGCAGC 4735
 QY 540 GlnGlnSerProGlyGlyAspPheProAlaMetThrAsnAsnValGlySerAsnMet 559
 Db 4736 CAGCAGTCGCGAGGTGGAGATTTCCCGCGCAATGACGAATAATGTTGGCTCTAATG 4795
 QY 560 TyrTyrHisGlyGlyGlyGlyGlyGlyValAlaProThrPheThrValTyrAsnAspAsn 579
 Db 4796 TATTACCATGGGAAGTGGTGGAGAGTTGCTCCACATTTACAGTTTGGACGACAAAT 4855
 RESULT 6
 AF317906 4873 bp DNA linear PLN 10-OCT-2002
 LOCUS Brassica napus AP2/EREBP transcription factor BABY BOOM1 (BBM1)
 DEFINITION Gene, complete cds.
 ACCESSION AF317906
 VERSION AF317906.1 GI:21069054
 KEYWORDS
 SOURCE Brassica napus (rape)
 ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 4873)
 Boutillier,K., Offringa,R., Sharma,V.K., Kieft,H., van
 Lammeren,A.M., Quillet,T., Zhang,L., Hattori,J., Liu,C.-M.,
 Mixi,B.L.A., Custers,J.B.M. and van Lookeren Campagne,M.M.
 Ectopic expression of BABY BOOM triggers a conversion from
 vegetative to embryonic growth

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Plant Cell 14 (8), 1737-1749 (2002)
 22163009
 12172019
 2 (bases 1 to 4873)
 Boutillier,K. and Fiers,M.
 Direct Submission
 Submitted (02-NOV-2000) Business Unit Plant Development and
 Reproduction, Plant Research International, Droevendaalsesteeg 1,
 Wageningen 6708 PB, The Netherlands

FEATURES
 source

1. .4873
 /organism="Brassica napus"
 /mol_type="genomic DNA"
 /cultivar="Topas"
 /db_xref="taxon:3708"
 <1620. .>4858
 /gene="BBM1"
 join(<1620. 1845,2299. 2719,2953. 3035,3161. 3169,
 3315. 3403,3554. 3627,3798. 3848,3962. 4038,4149. .>4858)
 /gene="BBM1"
 /product="AP2/EREBP transcription factor BABY BOOM1"
 join(1620. 1845,2299. 2719,2953. 3035,3161. 3169,
 3315. 3403,3554. 3627,3798. 3848,3962. 4038,4149. .4858)
 /gene="BBM1"
 /codon_start=1
 /product="AP2/EREBP transcription factor BABY BOOM1"
 /protein_id="AA033802.1"
 /db_xref="GI:21069055"

gene

mRNA

CDS

Alignment Scores:

Pred. No.: 2,616-183 Length: 4873
 Score: 2690.50 Matches: 574
 Percent Similarity: 53.24% Conservative: 1
 Best Local Similarity: 53.18% Mismatches: 4
 Query Match: 86.28% Indels: 505
 DB: 8 Gaps: 8

US-09-980-364-2 (1-579) x AF317906 (1-4873)

QY 1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisSerArgLys 20
 Db 1620 ATGATATATTAAGTGTAGGCTTTCTCTCTCTTAAGAACAAATCACATCGTAAG 1679
 QY 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
 Db 1680 GAGCTCTACTCTTCCACCAACCGTCGTAGATGTCGCCGAGAGTACTGTTCAGAT 1739
 QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
 Db 1740 CCGACCGCTGCTCGATGAGTCTTCAGCCATCAAAACATCGTTTCTCTCCCTTGGT 1799
 QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArg----- 75
 Db 1800 GTCTGCTGTCGATGCTTTCCACGAGACAAACAATAGTCACTCCCGAGGTTATTGTTTGA 1859
 QY 75 ----- 75
 Db 1860 ACTACTTCTTTTTTTTGGTTTATTTGTTAGTTTCTCTCTTCTTCCATCGTAGA 1919
 QY 75 ----- 75

ORIGIN


```

QY 75 ----- 75
Db 28880 CATTTTCGTATATTATAGATCGGTGATGAATCAAAACCCACCGTTGGGGATCGCTTTGTTTT 28821

QY 76 ----- AspTrpAspIleasnGlyCysAlaCysAsnAsnIleHisAsnAspGluGln 92
Db 28820 TTGTCCTATAGATTGGGACATCAATGGTGGTCATGCAATAACATTAACAATAACGAACAA 28761

QY 93 AspGlyProLysLeuGluAsnPheLeuGlyArgThrThrIleTyAsnThrAsnGlu 112
Db 28760 AATGACCAAAAGCTTGAGAAATTCCTGGCGCCACCCACCGATTACAAATACCAACGAG 28701

QY 113 AsnValGlyAspGlySerGlySerGlyCysTyArgGlyGlyAspGlyGlyGlySer 132
Db 28700 ACCGTTGTAGATGGAAATGGCGAT---TGTCGAGGAGGAGACGGTGGTGGCGCGCTCA 28644

QY 133 LeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsnValAspAsn 152
Db 28643 CTAGGCCCTTCGATGATAAACATGGCTGAGTAATCATTCGTTGCTAATGCTAATCAT 28584

QY 153 GlnGluAsnGlyAsnAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCys 172
Db 28583 CAAGCAATGGTAAACGGTGGCAGAGGCTTGTCCCTCTCTATGAATTCATCTACTAGT--- 28527

QY 173 AspAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThrIleAspAspSer 192
Db 28526 GATAGCAACACTACACAAACATGATGATGCTGCTCCAGAGAGAGACTATTGTTGATGTC 28467

QY 193 ValGluAlaThrProLysThrIleGluSerPheGlyGlnA-gThrSerIleTyArg 212
Db 28466 GTAGAAACTACACCGAAGAAACTATTGAGAGTTTGGACAAAGCAAGCTCTATATACCGC 28407

QY 213 GlyValThr----- 215
Db 28406 GGTGTTTAC-AAGGTTAAATTCATTCATCTATGTATATATTTTATGTCGTTAAATGTGAT 28348

QY 215 ----- 215
Db 28347 TTTCCTGGTATTGTTGGACATTCCTAATGGTTCGGTTGAGAGAGAGTGCAACGGAATGT 28288

QY 215 ----- 215
Db 28287 CTCTCAATGTATATTAAGAGAAACATTAATAGTGATACATGGGTTTATATATACAATAA 28228

QY 216 ----- Arg 216
Db 28227 TACGTCATATATATGGTATGCTCTGATCATAGTATATATATGTTGCTTAAATGTGAT 28168

QY 217 HisArgTrpThrGlyArgTyGluAlaHisLeuTrpAspAsnSerCysLysArgGlyGly 236
Db 28167 CATCGGTGCAGAGTAGATACAGGCAATTTATGGGACATAGTTGCCAAAAGAGAAGGC 28108

QY 237 GlnThrArgLysGlyValGlnValTy----- 245
Db 28107 CAGACTCCAAAGAGAGACAGG-TACTATATATAAAGCTAATTTTAAATTTTCATT 28049

QY 245 ----- 245
Db 28048 TACCATTATTTTCAAACTAATTTAGTGTTCCTTTTCATGTGTTTCATCAAAATTTGCAC 27989

QY 245 ----- 245
Db 27988 CTGATGGCTCTCTTTTCAAGTTTATCTCGGTAAGTTCCTTGATTTTAAGCTATATAATTAATA 27929

QY 246 ----- Leu 246
Db 27928 ATAGATGACTATTAAATCTATTCTAAGCAAAATATAATTTGTTGTTATCTGATCCTTACA 27869

QY 247 GlyGlyTyAspLysGluGlyAlaAlaArgAlaTyAspLeuAlaAlaLeuLysTy 266
Db 27868 GGAGTTTATGAAGAAGAAAGAGAGCTAGGCTTACGATTTAGCCGCACTAAAGTAT 27809

```

```

QY 267 TrpGlyThrThrThrThrThrAsnPhe-Pro----- 276
Db 27808 TGGGGAACCAACCACTACTACTAACTTCCCGTATGTTAATTAATCAATAATATATACATA 27749

QY 276 ----- 276
Db 27748 AATTCCTAACTTCTAACCAATTTTAGTCTGAATAATGCCAATCTCTTAACTAGTATTAT 27689

QY 276 ----- 276
Db 27688 CTTACTATTAACTGTCATGCTTTATATTGTTACATTTGTTACAAATAAAATAGTAATGTT 27629

QY 277 ----- MetSerClyTyArgLysGluValGluGluMetLysHisGly 290
Db 27628 GGTGGATATAATATTCAGTTGAGTGAATATAGAAAGAGGTAGAGAGATGAACACAT 27569

QY 290 tThrArgGlnGluTyValAlaSerLeuArgArg----- 301
Db 27568 GACGAGGCAAGAGTATGTTGCTCTCTGCGCAGGTACAGAAATGAAACTCTTTGAATTTATT 27509

QY 301 ----- 301
Db 27508 GCATTTTAGAAACCCCATCAGCTATATATTTATTAATAATATATCGTAACATTGAATAATC 27449

QY 302 ----- LysSerSerGlyPheSerArg 308
Db 27448 ATTATTTGGAAGATATAAGAACAACATGTAATATATCGAGGAAAGT-AGTGGTTTCTCTCG 27390

QY 308 gGlyAlaSerIleTyArgGlyValThr----- 317
Db 27389 TGGTGCATCGATTATTCGAGGAGTAAACAAGTACGTATAATCCATCTAGATATGGAACGA 27330

QY 317 ----- 317
Db 27329 ATACAGTGTTCATTATTTTTCATGTATACATAATAATTTGTCATACATACTATTAA 27270

QY 318 ----- ArgHisHisGlnHisGlyArgT 325
Db 27269 TCTAATCTAATAATATTTTCCTTTTAAATGTTTCCAAAAGGCATCACCACATGGAAGGT 27210

QY 325 rGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyLeuGlyThrPhe----- 343
Db 27209 GCAAGCTAGGATCGAAGAGTCGCGGTAAACAAGACCTCTACTTGGAACTTTCGTA 27150

QY 343 ----- 343
Db 27149 CATTTTCCAATAAATCTATATCTATTAAGATATTAAATATATACAAATATATCTAAGTG 27090

QY 344 ----- GlyThrGlnGluAlaAlaGluAlaTyAspIleAla 357
Db 27089 AATCATACAAATATTGTAGGCACACAGGAAGAGGTGCTGAGGCTTATGACATTCGACCC 27030

QY 358 IleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyAsnValLys 377
Db 27029 ATTAAATTCAGAGGATTAAAGCCAGTGAATACTTCGACATGAACAGATACATGTTAA 26970

QY 378 AlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAla 397
Db 26969 GCAATCCTCGAGAGCCGAGTCTACCTATTGTTAGTCTTTCGAAACGCTCTCAAGGACGTT 26910

QY 398 AsnArgProValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSer 417
Db 26909 AATAATCCGGTTCAGCT---ATGATGATTAGTAATAACGTTTCAGAGAGTGCATAATA 26853

QY 418 AlaSerGlyTrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeu 437
Db 26852 GTAGCGGTGCAAAACACTGCGTTTCAGCATCATCAGGAATGAGATTGAGCTTATTG 26793

QY 438 HisGlnHisGlnGluArgTyAsnGlyTyTrpTyAsnGlyGlyAsnLeuSerSerGlu 457
Db 26792 CAGCAACAGCAGGAGGATCGTTGGT---TATTACATGAGGAACTTCTTACCGAG 26736

QY 458 SerAlaAla-gAlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGln 477

```


Db 3042 TTCTTGGTATTGTTGGGACATTCCTAATGGTTCGGTTGAGAGAGAGTGCAACGGAATGT 3101
 QY 215 ----- 215
 Db 3102 CTCTCAATGATATTAAGAGAGAAACATTAATAGTGTACATGGGTTTATATATACAATAA 3161
 QY 216 -----Arg 216
 Db 3162 TACGTCATATATATCGTATGCTCTTGATCATAGTATATAATAGTTTGAATTTAATGTGAGG 3221
 QY 217 HisArgTTPThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGly 236
 Db 3222 CATCGGTGACAGGTAGATACGAGGCACATTTATGGGACAATAGTTGCAAAAGAGAAGGC 3281
 QY 237 GlnThrArgLysGlyArgGlnValTyr----- 245
 Db 3282 CAGACTCGCAAGGAAGACAAGG-TACTATATATATAAAGCTAATTTTAAATTTTCATT 3340
 QY 245 ----- 245
 Db 3341 TACCAATTTATTTCAAACTAATTTAGGTTTCTTTTCATGTGTTTCATCAAAATTTGCAC 3400
 QY 245 ----- 245
 Db 3401 CTGATGGCTCTCTTTTCAGTTTATCTGGGTAAGTCTTGTATTTAAGCTATAAAATTAATA 3460
 QY 246 -----Leu 246
 Db 3461 ATAGATGACTATTAAATCTATTCAGCAAAATATATAGTTGTGTATCTGATCCTTACA 3520
 QY 247 GlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyr 266
 Db 3521 CGAGGTTATGACAAAGAGAAAAGCAGCTAGGCTTACGATTTAGCGGCATTAAGTAT 3580
 QY 267 TrpGlyThrThrThrThrAsnPhe-Pro----- 276
 Db 3581 TGGGAGCCACCACTACTACTAACTTCCCGTATGTTAATTAATCAATAATATATACATA 3640
 QY 276 ----- 276
 Db 3641 AATTCCTAACTCTTAACCAATTTTAGTCTGAATAATGCCAATCTCTTAAACTAGTATTAT 3700
 QY 276 ----- 276
 Db 3701 CTTACTATTAACTGTCATGTTTATATGTTTACAATAAAATAGTAATGTTGGTGATA 3760
 QY 277 -----MetSerGluTyrGluLysGluValGluMetLysHisMetThrArgG1 293
 Db 3761 TAATATTTCAGTTGAGTGAATATGAAAGAGGTAGAGAGATGAACACATGACGAGGCA 3820
 QY 293 nGluTyrValAlaSerLeuArg----- 301
 Db 3821 AGAGTATGTTGCCTCTCTGCGCAGGTACAGAACTCTTGAATTTATTGCAATTTTAG 3880
 QY 301 ----- 301
 Db 3881 AAACCCATCACGTATATATTATTAATAATATATCGTAACATTCGTAATTAATTAATTTGG 3940
 QY 302 -----LysSerSerGlyPheSerArgGlyAlaLase 311
 Db 3941 AAAGATATAAGAAACATGTAATATATGACGAGAAAAGT-AGTGGTTTCTCTCGTGTGTCATC 3999
 QY 311 rIleTyrArgGlyValThr----- 317
 Db 4000 GATTTATCGAGAGTAAAC-AAGGTACGTATATATCCATCTAGATATGAAACGAATACTAGT 4058
 QY 317 ----- 317
 Db 4059 GTTTCATTATTTTGTGATATACATAATAATATGTCATACATACTACTATTAATCTAATC 4118
 QY 318 -----ArgHisGlnHisGlyArgTTPGlnAl 327
 Db 4119 TAATTAATATTTCCTTTAAATGGTTCAAAAGGCAATCCAAACATGGAAGGTGGCAAGC 4178

QY 327 aArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPhe----- 343
 Db 4179 TAGGATCGGAAGAGTGCOCGGTAAACAAAGACCTCTACTTGGGAACCTTTTCGGTACATTTTC 4238
 QY 343 ----- 343
 Db 4239 CAATAAAATCTATATACATATAAGATATTAAATATATACAAAATATATCTAAGTGAATCATA 4298
 QY 344 -----GlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaLysP 360
 Db 4299 CAAATTTATGTAGGCACACAGGAAGAGGCTGCTGAGGCTTATGACATTCAGCAATTAAT 4358
 QY 360 heArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleL 380
 Db 4359 TAGAGGATTAGCGCAGTACTAACTTCGACATGACAGATACATATGTTAAAGCAATCC 4418
 QY 380 euGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgP 400
 Db 4419 TCGAGAGCCGAGTCTACCTATTGCTAGTTCTCGAAACGTCCTCAAGGACGTTAACAAATC 4478
 QY 400 roValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerG 420
 Db 4479 CGGTTCCAGCT---ATGATGATTAGTAAATAGCTTCAGAGAGTCCAAATTAATGTTAGCG 4535
 QY 420 lyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGlnH 440
 Db 4536 GTTGGCAAAACACTCGGTTTCAGCATCATCAGGGAATGGATTGAGCTTATTTCAGCAAC 4595
 QY 440 isGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaA 460
 Db 4596 ACACAGAGAGTACCTGGT---TATTACAATGGAGAAACTTTGTCTACCGAGAGTACTA 4652
 QY 460 rGAlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuM 480
 Db 4653 GGGTTTGTTCAAACAGAGGAGGAAACAACACTCTCTTGAGAAACTCGCGAGTCA 4712
 QY 480 eThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnV 500
 Db 4713 TGACTAATGTTGATCATCATGCTCGACCTCTGATGATTTCTGTTACCGTTTGTGGAAATG 4772
 QY 500 alValGlyTyrGlyTyrGlnGlyPheAlaAlaPro-----ValAsnCysA 516
 Db 4773 TTGTTAGTTATGCTGTTATCAAGATTCGCAATCCCTGTTGGAACATCGGTTAATTACG 4832
 QY 516 spAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAla---- 534
 Db 4833 ATCCCTTTACTGCTGCTGAGATTGCTTACAACGCAAGAAATCAATTATTACTATGCTCAGC 4892
 QY 535 -----GlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMetThrA 553
 Db 4893 ATCAGCAACACAGCAGATTCAAGAGTCCCGGAGGAGATTTTCGGTGGCGATTTCGA 4952
 QY 553 snAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaProThrP 573
 Db 4953 ATAACCATAGCTCTAATCATGTACTTTTCCGGGGAAGGTGGTGGAGAGGGGCTCCACGT 5012
 QY 573 heThrValTrpAsnAsp 578
 Db 5013 TTTCACTTTTGGAAACGAC 5029

RESULT 9

AF317907 5151 bp DNA linear PLN 10-OCT-2002
 LOCUS Arabidopsis thaliana BABY BOOM (BBM) gene, complete cds.
 DEFINITION Arabidopsis thaliana
 ACCESSION AF317907
 VERSION AF317907.1 GI:21069056
 KEYWORDS SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.

```

REFERENCE 1 (bases 1 to 5151)
AUTHORS Boutillier,K., Offings,R., Sharma,V.K., Kieft,H., van
          Lammeren,A.M., Ouellet,T., Zhang,L., Hattori,J., Liu,C.-M.,
          Miki,B.L.A., Custers,J.B.M. and van Lookeren Campagne,M.M.
TITLE    Ectopic expression of BABY BOOM triggers a conversion from
          vegetative to embryonic growth
JOURNAL  Plant Cell 14 (8), 1737-1749 (2002)
MEDLINE 22163009
PUBMED  12172019
REFERENCE 2 (bases 1 to 5151)
AUTHORS Boutillier,K. and Sharma,V.K.
TITLE    Direct Submission
JOURNAL  Submitted (02-NOV-2000) Business Unit Plant Development and
          Reproduction, Plant Research International, Droevendaalsesteeg 1,
          Wageningen 6708 PB, The Netherlands
FEATURES
  source Location/Qualifiers
    1. 5151
       /organism="Arabidopsis thaliana"
       /mol_type="genomic DNA"
       /isolate="C24"
       /db_xref="taxon:3702"
       /chromosome="v"
       <2017..>5035
       /gene="BBM"
       join(<2017..2248,2579..2993,3221..3303,3420..3428,
       3522..3610,3771..3844,3970..4020,4152..4228,4311..>5035)
       /gene="BBM"
       /product="BABY BOOM"
       join(2017..2248,2579..2993,3221..3303,3420..3428,
       3522..3610,3771..3844,3970..4020,4152..4228,4311..5035)
       /gene="BBM"
       /codon_start=1
       /evidence=not experimental
       /product="BABY BOOM"
       /protein_id="AM33803.1"
       /db_xref="GI:21069057"
       /translation="MNSMNLGSLSPHQNHRTDVSSTRTTAVDVAGGCFDLA
       APDSASVATSLSPGVLEAFTDNNSHSDWDINGACNLTNNQNGPKLENF
       LGRTTITYNRTVWDGDCGGDGGSLGLSMIKTWSHVSANAHQDNGNGA
       RGLSNSTSDSNNNNDVQKTIIVDVETTPKTIIESFGQRTSIRYGVTRHR
       WTRGSEAHLDNSCKRGOTKRGVYVLLGGYDKKAAAYDLAALKYWGPTTTNFP
       LSYEKVEVEMKNTQEVYASLRKSSGFSRGASLYRGVTRHHQGRQARIGRVAG
       NKDLYLGTGTQBEAEYDIAALKRGLSAVTNFDNRNINVKALIESLPIGSSAK
       RLKDNVTFPAMMISNNVSANNVGWQNTAFQHHQGMDSLQQQKERYVYGGQ
       NLSTESPRVCFKQEEQHFRLNSPHTMNVDSHSDSDSVTCNVSVSYGGYQQFA
       IPVGTSYNDPFFAAEIAIYARNHYHQAHQQQQQIQQSPGGDFPVAISNNHNNYF
       HGEGGGGGGAFTFSWMDT"
ORIGIN
Alignment Scores:
Pred. No.: 5,14e-139 Length: 5151
Score: 2073.50 Matches: 469
Percent Similarity: 50.15% Conservative: 36
Best Local Similarity: 46.57% Mismatches: 67
Query Match: 66.48% Indels: 439
DB: 8 Gaps: 13

US-09-980-364-2 (1-579) x AF317907 (1-5151)

Qy 3 AsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLysAspVal 22
Db 2029 AATAACTGGTTAGGCTTCTCTCTCTCTCTCATGATCAAAATCATCACCGTACCGAATGT 2088

Qy 23 TyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAspProThr 42
Db 2089 GACTCCTCCACCACAGAACCGCGTAGATGTTCCCGAGGGGTACTGTTTGTATCTGGCC 2148

Qy 43 AlaAlaSerAspGluSerAlaIleGlnThrSerPheProSerProPheGlyValVal 62
Db 2149 GCTCCCTCCGATGATCTCTCTGCGGTCCAAACATCTTTTCTTCTCTCTCTCGGTGTCACC 2208

Qy 63 ValAspAlaPheThrArgAspAsnAsnSerHisSerArg----- 75
       ::::|||||

```

```

Db 2209 CTCGAGGCTTTCACCAGAGACAATAAGTAGTCACTCCGAGGTTGTGTTTTAAAAATATT 2268
Qy 75 ----- 75

Db 2269 TATTTTATCTTTGTTTGTATTATTTTTCCTCTTCTTCCAATGCATAGACAAACACAA 2328
Qy 75 ----- 75

Db 2329 GACTCACGCACGTAGCCCTATTTTGTGTTTTCATTGTTTATCGATTTCATCTCTTTGAG 2388
Qy 75 ----- 75

Db 2389 AATTTCCATGAGTGGGTTTAGTGTGTTGTTGTCACATGATCACATCTCATGAATTTAAACT 2448
Qy 75 ----- 75

Db 2449 AGTAAACATGAACACTAGACATTTATTTTGTACCTTTTATCTTATATAAATGAATTC 2508
Qy 75 ----- 75

Db 2509 CATTTCTGTATATTATAGATCGGTGATGAATCAAAACCAACGTTGGGGATCGCTTTGTTT 2568
Qy 76 -----AspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGln 92
Db 2569 TTGTCCTATAGATTGGGACATCAATGTGTCGATGCAATACATTAACCAATACCAACAA 2628
Qy 93 AspGlyProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThrAsnGlu 112
       ::::|||||
Db 2629 AATGACCAACAGCTTGAGAAITTCCTCGCCGCCACCCACGATTTTACAATACCAACAG 2688
Qy 113 AsnValGlyAspGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGlySer 132
       ::::|||||
Db 2689 ACCGTTGTAGATGGAATGGCGAT---TGTGGAGGAGGAGACGGTGGTGGTGGCGGCTCA 2745
Qy 133 LeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsnValAspAsn 152
       ::::|||||
Db 2746 CTAGCCCTTTCATGATATAAACATCGCTGAGTATCATTCGTTGCTTAATGCTAATCAT 2805
Qy 153 GlnGluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCys 172
       ::::|||||
Db 2806 CAAGCAATGTTAACCGTGCACGAGCTTGTCCCTCTCTATGAATTCATCTACTAGT--- 2862
Qy 173 AspAsnAsnAspSerAspAsnAsnValValAlaGlnGlyLysThrIleAspAspSer 192
       ::::|||||
Db 2863 GATAGCAACATCTACAAACCAATGATGTCGTCGCAAGAGACATTTGTTGATGTC 2922
Qy 193 ValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArg 212
       ::::|||||
Db 2923 GTAGAACTACACCGAAGAAACTATTGAGAGTTTTGCAAAAGACGCTCTATATACCGC 2982
Qy 213 GlyValThr----- 215
       :|||
Db 2983 GGTGTAC-AAGGTTAATTTTCATGATCTATGATATATTTTATGCTTAATTTGTGAT 3041
Qy 215 ----- 215

Db 3042 TTTCTTGGTATTGTTTGGGACATTTCTAATGGTTCGTTGAGAGAGAGTCAACGGAATGT 3101
Qy 215 ----- 215

Db 3102 CTCTCAATGATATATAAAGAGAAACATTAAATAGTGTACATGGGTTTATATATACAATAA 3161
Qy 216 -----Arg 216
       :|||

Db 3162 TAGGTCATATATATGTTGCTCTTGTATCATAGTATATAATGTTTGAATTTAATGTCAGG 3221
Qy 217 HisArgTTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGly 236
       :|||
Db 3222 CATCGGTGACAGGTAGATACCGAGGCACATTTATGGGCAATAGTTGCAAAAGAGAGAGGC 3281
Qy 237 GlnThrArgLysGlyArgGlnValTyr----- 245
       :|||
Db 3282 CAGACTCGCAAGGAGACACAGG-TACTATATATATAAGCTAAATTTTAAATTTTCATT 3340
       :|||

```

QY 245 ----- 245
 Db 3341 TACCATTATTTTCAAACTAATTTAGGTTTCTTTTCATGTTGTTTCATCAAAATTTGCAC 3400
 QY 245 ----- 245
 Db 3401 CTGATGGCTCTCTTTTCAGTTTATCTGGGTAAGTCTTTGATTTTAAGCTATAAATAA 3460
 QY 246 -----Leu 246
 Db 3461 ATAGATGACTATTAATCTATTTCTAAGCAAAATATAATTTGTTGTTATCTGATCCTTACA 3520
 QY 247 GlyGlyTyrAspLysGluGluLysAlaAlaAlaAlaAlaAlaAlaAlaAlaLeuLysTyr 266
 Db 3521 GGAGGTTATGCAAAAGAAAGAAAGCAGCTAGGGCTTACGATTTAGCGCACTAAAGTAT 3580
 QY 267 TrpGlyThrThrThrAsnPhe-Pro----- 276
 Db 3581 TGGGACCCACCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 3640
 QY 276 ----- 276
 Db 3641 AATTCCTAACTCTTAACCAATTTAGTCTGAATAATGCAATCTCTTAAACTAGTATTAT 3700
 QY 276 ----- 276
 Db 3701 CTTACTATACTGCTGTTTATATTGTTTCAATAAATAATGATAATGTTGGTGGATA 3760
 QY 277 -----MetSerGluTyrGluLysGluValGluMetLysHisMetThrArgG1 293
 Db 3761 TAAATTCAGTTGAGTGAATATGAGAAAGAGTAGAGAGATGAAGCAGATGACAGGCA 3820
 QY 293 hGluTyrValAlaSerLeuA-Arg----- 301
 Db 3821 AGAGTATGTCCTCTCTGCGCAGGTACAGAATGAACTCTTGAATTTATTGCAATTTAG 3880
 QY 301 ----- 301
 Db 3881 AAACCCATCAGTATATTTTAAATATATATGTAACATTTGAATTAATTAATTTGG 3940
 QY 302 -----LysSerSerGlyPheSerArgGlyAlaSe 311
 Db 3941 AAAGATATAAGAAACATGTAATATATGCGAGGAAAGT-AGTGGTTTCTCTGCTGTCGATC 3999
 QY 311 rIleTyrArgGlyValThr----- 317
 Db 4000 GATTTATCGAGAGTAAC-AAGGTACGTATTAATCGATCTAGATATGGAACGATACTAGT 4058
 QY 317 ----- 317
 Db 4059 GTTTCATTATTTTTCATGATACATAATAATTTGTCATACAATACTATTAACTTAATC 4118
 QY 318 -----ArgHisHisGlnHisGlyArgTrpGlnAl 327
 Db 4119 TAATTAATATTTCTTTTAAATGGTTCCAAAGGATCACCACATGGAAGGTGGCAAGC 4178
 QY 327 aArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPhe----- 343
 Db 4179 TAGGATCGGAAGAGTCGCGGTACAAAGACCTCTACTTGGGAACCTTCGTCATCATTTTC 4238
 QY 343 ----- 343
 Db 4239 CAATAAAATCTATATCTATTAAGATATTAATATATACAAATATATCTAAGTGAATCATA 4298
 QY 344 -----GlyThrGlnGluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysP 360
 Db 4299 CAAATTAATGAGGCACAGAGAGGCTGCTGAGGCTTATGACATTCAGCCATTAAT 4358
 QY 360 hArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleL 380
 Db 4359 TCAGAGGATTAAGCGCAGTACTAACTTCGACATGAACAGATCAATGTTAAAGCATCC 4418

QY 380 euGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgP 400
 Db 4419 TCAGAGAGCCCGAGTCTACCTATTGTTAGTTCTCGAAACGCTCTCAAGGACGTTAAACAATC 4478
 QY 400 roValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerG 420
 Db 4479 CGGTTCCAGCT---ATGATGATTAGTAATTAAGCTTTCAGAGAGTCAAAATATGTTAGCG 4535
 QY 420 LYTPGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeuHisGlnH 440
 Db 4536 GTTGGCAAAACACTGGTTTCAGCATCATCAGGAAATGGAATTTAGCTTATTCGACCAAC 4595
 QY 440 isGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaA 460
 Db 4596 AGCAGAGAGAGTACGTTGGT---TATTACAAATGGAGAACTTGTCTACCGAGAGTACTA 4652
 QY 460 rGAlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeu 480
 Db 4653 GGGTTTGTTCACACAGAGGAGGACACACACTTCTTGAGAACTCGCGAGTCA 4712
 QY 480 euThrAsnIleAspHisGlnSerValSerAspSerValThrValCysGlyAsnV 500
 Db 4713 TGACTAATGTTGATCATCATAGCTCGACCTCTGATGATTCTGTACCGTTTGTGGAATG 4772
 QY 500 alValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaPro-----ValAsnCysA 516
 Db 4773 TTGTTAGTTATGTTGTTATCAAGGATTCGCAATCCCTGTTGGAACATCGGTTAATTACG 4832
 QY 516 sPalatyrAlaAlaSerGluPheAspTyrAsnAlaArAsnHisTyrTyrPheAla---- 534
 Db 4833 ATCCCTTTACTGCTGTGAGATGCTTACACCGCAAGAAATCATTTACTATGTCAGC 4892
 QY 535 -----GlnGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaKetThrA 553
 Db 4893 ATCAGCAACAAACAGCAGATTCAGCAGTCGCGGAGGAGATTTTCGGTGGCATTTTCA 4952
 QY 553 snAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaProThrP 573
 Db 4953 ATAACCATGCTTAACATGTAATTTTACGCGGGAAGGTGTTGGAGAGGGCTCCACAGT 5012
 QY 573 heThrValTrpAsnAsp 578
 Db 5013 TTTCAGTTTGGAAACGAC 5029
 RESULT 10
 AY461432
 LOCUS Nicotiana tabacum ANT-like protein mRNA, complete cds.
 DEFINITION AY461432
 ACCESSION AY461432.1 GI:38492171
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Nicotiana tabacum (common tobacco)
 Nicotiana tabacum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
 asterids; Lamiales; Solanales; Solanaceae; Nicotiana.
 1 (bases 1 to 2546)
 Rieu, I. and Mariani, T.
 A tobacco ANT homolog is expressed in the whole ovary
 Unpublished
 2 (bases 1 to 2546)
 Rieu, I. and Mariani, T.
 Direct Submission
 Submitted (07-NOV-2003) Plant Cell Biology, University of Nijmegen,
 Toernooiveld 1, Nijmegen 6525 ED, The Netherlands
 FEATURES
 1..2546
 Location/Qualifiers
 /organism="Nicotiana tabacum"
 /mol_type="mRNA"
 /cultivar="petit havana SR1"
 /db_xref="taxon:4097"
 343..2274
 /note="similar to Arabidopsis thaliana ANT protein"
 CDS

```

/codon_start=1
/product="ANT-like protein"
/protein_id="AAR22388.1"
/db_xref="GI:38492172"
/translation="MKKSNDDNNCSHRSNNNNNNNSNDASNLGFLSLP
HKWETSSEPHQHQPNNQSSALAQSYLSSPPNNVTSNLSALCYGVGENNPPGHS
DSLVMLPKSDGLICINEARSHADAWQSSPKLEDFLGASQVGNHERAMALS
DSLYHONEGYYSPMHCHMYQESLDETKPTQISHCDAQMTAVSGNELKSGHVAID
QHINETCSMVAGGCGGSGGSGTGGNDLQSLSLSMNPGSQSCVTPRQISPNEL
ECVAITKRAQKVQKQPVHRKSIDTFQRTSOVYTRHRWTRVEAHLWDSCK
KEGCTKQKVLYGGVDMBEKAAKYDLAALKYWPSTHINFLPNYOKELMDKQMT
ROYVAHLRKSSGFSRGSAMVYGVTRHHQHGWRQARIGRVAGNKLVLSTPSEEA
AAAYDVAALKFRGVNAVTHFDISRDYVEXIMASNTLIPAGELARRNKERPIEYNLSDH
KNEEACVQNNNNNNNDWQVSYQSNFSLGNRPFTSMALQDLIGLIDWNSS
NHILDDQQKIGNHFNSASLVTSLGSSREASPKSAASLVFAKPTNVNACIPSAQL
RPIPVSHLPVFAALNDA"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      6,02e-63      Length:      2546
Score:          1008.00      Matches:      258
Percent Similarity: 50.08%      Conservative: 65
Best Local Similarity: 40.00%      Mismatches: 132
Query Match:      32.32%      Indels:      190
DB:              8          Gaps:      24

```

US-09-980-364-2 (1-579) x AY461432 (1-2546)

```

Qy 2 AsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLysAsp 21
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 442 AGTAGTAACCTGGTAGGATTTCTCTCTCACCC-----CACATGAAATGGAG 489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 22 ValTyrSerSerThrThrThrValValAspValAlaGlyLutyrCysTyrAspPro 41
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 GTTACTTCTTCT-----GAACCTCAACACCAACAT 519
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 42 ThrAlaAlaSerAspGluSerAlaIleGlnThrSerPhe-----ProSerPro 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 520 CATCAGTTCAATAATCAGAGCTCTGCTCTTGGCTCAAGTTTCTATCTTCTCTCTCT 579
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 59 PheGlyValValValAlaPhe-----ThrArgAspAsnAsn----- 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 580 ATGAATGTGACTTCAATACTCTGCTCTGTGTATGTTGGAGAGATAATCCCTTT 639
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 72 SerHisser-----ArgAspTrpAspIleAsnGlyCysAlaCys----- 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 640 GGCAATCTCTTTATCTGTATGCTCTCAAAATCTGATGCTCTTTCATATATGGAA 699
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 85 -----AsnAsnIleHisAsnAsp-----GluGlnAspGlyProLysLeuGlu 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 700 GCTCTCTCAAGATCTCATGAGATGCTATGCTGCAAAAGTTCTTCTTCAACCTTGG 759
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 99 AsnPheLeuGlyArgThr-----ArgAspTrpAspIleAsnGlyCysAlaCys 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 760 GACTTTTATGTTGGTCAACTCAATATGGAATCATGAAGGAAAGCTATGGCTCTTAAGC 819
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 105 -----ThrThrIleTyrAsnThrAsnGluAsn----- 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 820 TTAGACAGCTTATATTATCACCRAAATGAAGCTATTATTCTCCATGCATTGTCATGCC 879
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 113 ----- 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 980 ATGTATCAAGAAATCATTATTAGATGAACAAACAACTCAAAATTTCAATTTGATGCC 939
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 113 ----- 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 940 CAAATGACTGCAGTCTCTGGAATGAACCTGAAAGTTGGGGCATTTATGCTATTGACCAA 999
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 114 -----ValGlyAspGlySerGlySerGlyCysTyrGly 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1000 CATATAAATGAACCTTGTAGTAGTATGTTGGTGCAGGTGGTGGAGGTACTTCTCGT 1059
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 125 GlyGlyAspGlyGlyGlyGlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsn 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1060 -----TCTGTGGTGGTACAGTTGGT----- 1080
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 145 GlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLysGlyLeuSerLeu 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1081 -----GGTAATGATTTACAGTCTTTTGAGCTTG 1107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 165 SerMetAsn-----SerSerThrSerCysAspAsnAsnAsnAspSerAsnAsnVal 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1108 TCTATGAATCTCGTTCTCAGTCTAGCTGTGTTACTCCAAGACAAATTTCTCTCAATAGAG 1167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 183 Val-----AlaGlnGlyLysThrIleAspAspSer 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1168 CTGGAATCTGTTCATTAAGAAAGGAGGCTTCTGAAAGATGTTGTTCAAAA--- 1224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 193 ValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArg 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1225 ---CAACCTGTTTCATAGGAAATCTATTGATACATTTGGTCAGAGAACTTCTCAGTATAGA 1281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 213 GlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCys 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1282 GGTGTCTACTAGGCATAGGTGACAGGTAGGTATGAGCGCATTTTGTGGGATTAATAGTTGC 1341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 233 LysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGlu 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1342 AAGAAGGAAGGCGACACTAGAAAAGGAGGCAAGTTTATCTTTGGGGCTATGATATGGA 1401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 253 GluLysAlaAlaArgAlaTyrAspLeuAlaLysLysTyrTrpGlyThrThrThr 272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1402 GAGAAAGCTGCAAGAGCATATGATTTAGTCTACATAAGTATTGGGACCTTCAACTCAC 1461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 273 ThrAsnPheProMetSerGluTyrGluLysGluValGluMetLysHisMetThrArg 292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1462 ATTAATCTTTCGTTAGAGAACTACCAGAAAGAGCTTGAGGATATGAAGATATGACCCGG 1521
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 293 GlnGluTyrValAlaSerLeuArgLysSerSerGlyPheSerArgGlyAlaSerIle 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1522 CAAGAATATGTTGCACACTTAAAGGAAAGTAGTGGATTCTCAAGAGGTGCTTCAATG 1581
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 313 TyrArgGlyValThrArgHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgVal 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1582 TATAGAGAGTGACACAGCACCCATCAGCTGAGAGTGGCAGGCAAGATTTGGAAGAGT 1641
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 333 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluAlaGluAla 352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1642 GCAGGGAACAAAGATCTCTATCTTGGGACTTTCAGCACACAGAAAGAGCTGCGAAGCA 1701
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 353 TyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsn 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1702 TATGATGTTGTGCAATCAAAATTCAGAGGTGTAATGCTGTCAACAACTTTGACATATCG 1761
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 373 ArgTyrAsnValLysAlaIleLeuGluSerProSerLeuProIleGlySerAlaLys 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1762 CGATACGAGTGGAGAAATCATGGCTAGTAATACACTTCTCTGCTGGAGAAATTAGTAGG 1821
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 393 ArgLeuLysGluAlaAsnArgProValProSerMetMetIleSerAsnValSer 412
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1822 AGAAACAAAGAAAGAGAG-----CCAAATGAATACAACTCTCAGT 1860
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 413 GluSerGluAsnSerAlaSerGlyTyrGlnAsnAlaAlaValGlnHisGlnGlyVal 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1861 GATCACAGAATGAAGAGCATGTTTCAATGACAAAC---AACACAACAATGGGAAT 1917
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 433 AspLeuSer-----LeuLeuHisGlnHisGlnGluArgTyrAsnGlyTyrTyr 448
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1918 AATGTCACAGATTGGAAAATGTTGTCATATCAATCTCTCAAAATCTCTCTAGGAATAT 1977
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 449 TyrAsn-----GlyGlyAsnLeuSerSerGlu 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1978 AGAALCCCAACTTTTTCATGTCATTACAGATCTAATCGGTATTGATTGATGATTC 2037
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 458 SerAlaArgAlaCysPheLysGlnGluAspGln-----HisHisPhe 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db	1335	TGGCAATGAGGCATGTCAAATTGCTACCAACTGTCTTACATGACGACGACGA	1394
		:::	::: :::
Qy	467	paspGlnHisPheLeuSerAsnThrGlnSerLeuMetThrAsnIleAaspHisGlnSe	487
Db	1395	ACATCAGCACCATGACGACGACGACGACAA--	GATCAACAGGA 1439
Qy	487	rserValserAaspSerValThrValCysGlyAsnValValGlyTyrclyGlyTyrgl	507
Db	1440	TGATCAGTCACAAGCTCCAACACAGCTCGGCTCCAT--	1479
Qy	507	nGlyPheAlaAlaProVal-	513
Db	1480	-CGTTTGCAACACCACCTTCAGTTTCAGTGGTGCATATGAAGCTCCATGACGACGACG	1538
Qy	514	--AsnCysAspAlaTyraAlaIaserGluPhe-AspTyrAsnA	527
Db	1539	CACCTTTGGATACTACCAAAATGTTTCAGCTTTTCAGACACCAATCTTTGGGATGGA	1598
Qy	527	laArgAsnHisTyrrPheAlaGlnGlnGlnThrGlnGlnSerProGlyGlyAspp	547
Db	1599	ATTTCGGGAGGTGGCACTTCGACGCCAAAAGGATGAAGACCTGATCCTGATCAGAG	1658
Qy	547	heProAlaLaMetThrAsnAsnValGlySerAsn	558
Db	1659	TGATGAGGATCAAGACTGCATAGTTCATTCTCAC	1693
RESULT 12			
AX555220		1926 bp	DNA linear PAT 27-NOV-2002
LOCUS	AX555220	Sequence 5 from Patent WO02059332.	
DEFINITION	AX555220		
ACCESSION	AX555220	GI:259898748	
VERSION	AX555220.1		
KEYWORDS			
SOURCE	Oryza sativa		
ORGANISM	Oryza sativa		
REFERENCE	He,S. and Dotson,S.B.		
AUTHORS	Nucleic acid molecules associated with plant cell proliferation and		
TITLE	growth and uses thereof		
JOURNAL	Patent: WO 02059332-A 5 01-AUG-2002;		
FEATURES	Monsanto Technology LLC (US)		
source	Location/Qualifiers		
	1..1926		
	/organism="Oryza sativa"		
	/mol_type="unassigned DNA"		
	_db_xref="taxon:4530"		
	1..>1926		
	/note="unnamed protein product"		
CDS	/codon_start=1		
	/protein_id="CAD58227.1"		
	_db_xref="GI:259898749"		
	_db_xref="RENTREMBL,CAD58227"		
	/translation="MASGSGSSNWGLFSLSPHMAPVEVPSSSRSTAAHHHHHPAPPAH		
	AAAAAGASPPSATCTCNFLFPFAAQVAWPSGYXYGVGAVDGTSTAGVYVSHLE		
	VMPKIDGSUCITMEGMPPSPPLEDFLCGGNGSDHPDAIYYSCQRAEDASRAYOH		
	HQLPVNYQPLTAEMLEAAAAAMEDAMAAKNFLVTSYGACYGNQMCPQLSLMSNM		
	PQSQSSCVSAAPQQHQMAVAAAAGAAGDGQNSNDGGEQRVGKRGRTKGQKKOPF		
	VHRKSIDTFQRTSQVGRVTRHWRTGRYEHLWDNSCKDKQDTKRGKVYLGGVDTE		
	KAAARYDLAALYKGLSTHFNFLENRYDEIEEMERTROYVAHLRRRSRSGFSRAS		
	IYGVNTEHHOGHSQWARIIGSVAGNKDILCLGTFSTOEBAEAYDIATLKFGNLAVTNFN		
	DITRVDDTKMESSLLPGBAARKVKAIADPHPIGRELGAEASAATVCTDWR		
	MVLHGQQQAALACTATADLQGFMDGHASHLGIIVGFDVESNAADEIDVPQGKISG		
	INFNSGSSLVTSLSNRREGSPERLGLAMLAKKHHTFVSLAAMNPMMFMFAAAHVH		
	RPPSAIHLFPVFAAWTDA"		
ORIGIN			
Alignment Scores:	1.28e-60	Length:	1926
Pred. No.:	973-150	Matches:	254
Score:			

Alignment Scores:	1.28e-60	Length:	1926
Pred. No.:	973.50	Matches:	254
Score:			

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of
Agricultural Sciences Rice Full-Length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Coka, H., Hotta, I.,
Kojima, K., Naniki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Ohtomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
12869764

REFERENCE AUTHORS

2 (bases 1 to 2510)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T.,
Hori, F., Hotta, I., Ikeda, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koye, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Naniki, T., Nakikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Ohtomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, W., Tagami-Takeda, Y.,
Tagami, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.

TITLE JOURNAL

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agricultural Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp)
Tel: 81-29-838-7007, Fax: 81-29-838-7007
This clone is one of the 28K full-length cDNA clones from japonica
rice.

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cdna/>
NIAS Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Naniki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group; Ohtomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,

Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tanaka, T., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
source
1..2510
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiivar="Nipponbare"
/db_xref="taxon:39947"
/clone="002-101-C10"

ORIGIN

Alignment Scores:
Pred. No.: 1.77e-60 Length: 2510
Score: 973.50 Matches: 254
Percent Similarity: 50.48% Conservative: 63
Best Local Similarity: 40.45% Mismatches: 155
Query Match: 31.21% Indels: 158
DB: 8 Gaps: 19

US-09-980-364-2 (1-579) x AK106306 (1-2510)

QY 2 AsnAsnAsnTrpLeuGlyPheSerLeuSerPro----- 12
Db 200 AGCAGCAACTGGTAGGCTTCTCGCTCTCCCGCACATGCGGCGCATGGAGTGGCGCTCC 259
QY 13 -----TyGluGlnAsnHisArgLysAspValTyrSer 24
Db 260 TCCTCTGAGCATGACTGGCTGCTCATCATCATCATCATCATCATCATCATCATCATCATCTGCT 319
QY 25 SerThrThrThrValValAspValAlaGlyGluTyrCysTyrAspProThrAlaAla 44
Db 320 GCT-----GCTCGCGGAGCCATGCTGCTCTCTCCGACAGCGCC 358
QY 45 SerAsp-----GluSerSerAlaTleGlnThrSerPheProSerProPhe 59
Db 359 AGCAGCTCAACT 418
QY 60 GlyValValValAlaPhePheThrArgAspAsnAsnSerHisSerArgAspTrp----- 77
Db 419 TACTACTACGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 478
QY 78 -----AspIleasnGlyCysAlaCysAsnAsnIleHisAsnAsp 90
Db 479 CACCT 538
QY 91 GluGlnAspGlyProLysLeuGluAsnPheLeuGly-----Arg 103
Db 539 CGTCTCATCGCCCAAGCTCGAGGACTTCTTGGGTGTGGCAATGGCAGTGGCGGCGCATGAC 598
QY 104 Thr 119
Db 599 CGGGCCACCTACTATAGCCAGGCGGCAAGAGCAGCAGGATGCAAGCAG-GGGGGCGCTACCA 657
QY 119 ----- 119
Db 658 GCACCCAGCAGTAGTCCCTTCACTACCACTACCAGCCATTCACGGAAGCAGAGATGCTGCAAGA 717
QY 120 -----SerGlyCysTyrGlyGlyValAspGlyGlyGlyGlySerLeuGlyLeuSerMet 137
Db 718 GCGCGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 777
QY 138 IleTyrThrThrLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsn 157
Db 778 CTACGGCGCTCTGCTACGG-----CAACCA-----GGAGAT 807
QY 158 AlaAlaLysGlyLeuSerLeuSerMet-----AsnSerSerThrSerCys----- 172

QY 95 ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnVal 114
 Db 221 CCGAAACTGGAGACTTCCTC
 QY 115 GlyAspGlySerGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGlyGlySerLeuGly 134
 Db 242 -----GGCGCGCGCTGCACACGGGGGATCAAGCGTGGCGCTTG-CCCGCGGT 288
 QY 135 LeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGlu 154
 Db 289 -----CCAGACGACGGCGCGCCGACGCGCGCGCA 315
 QY 155 Asn----- 155
 Db 316 GCTGTACGAGTCGAGACTCAAGTTCTCGCGCGCGCTTCAGCTGAGCGCGCGCTGG 375
 QY 156 GlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsn 175
 Db 376 GGCGCGCGCGCGCTGCGCGCTGTTGCGCGCGCTGCTCTAGACCA-----GACCGA 429
 QY 176 AsnAspSerAsnAsnAsnValAlaGlnGlyLysThrIleAspAspSerValGluAla 195
 Db 430 CGAGACGAAGCAGCTGCGCTTGCCTGCGCGCGCGCGC-----CGTGGCGCGCGCGC 480
 QY 196 -ThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValTh 215
 Db 481 GGAGCAGAAGAGCGGTGACTCGTTCCGGCAGAGACGCTCTATCTACCGCGCGCTCAC 540
 QY 215 rArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgG 235
 Db 541 ACGGATCGGTGGACCGAAGTACGAGCGGCATCTGTGGGTAACAGCTGCGCGCGTGA 600
 QY 235 uGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGlyLysAl 255
 Db 601 GGCGCAGAGCGCAAGGACGCCAA-----GGTGGCTATGATAGGAGAGGAAGC 651
 QY 255 aAlaArgAlaTyrAspLeuAlaLeuLysTyrTrpGlyThrThrThrThrThrAsnPh 275
 Db 652 GCGAGGCGGTATGATCTTCGCGCTTGAAGTACTGGGTCTTAGCACCACCAACCT 711
 QY 275 eProMetSerGluTyrGluLysGluValGluGluMetLysMetLysHisMetLysArgGlnGluTy 295
 Db 712 TCCGGTTCCGAGTACGAAGAGGCTTGAGAGATGAAGCATGATGACGCCAGGAGTT 771
 QY 295 rValAlaSerLeuArgArgLysSerGlyPheSerArgGlyValaSerIleTyrArgG 315
 Db 772 TGTAGCTTCACTTAGGAGAGACAGCTGGATTCTCTCGTGGTCTCTATATACAGGG 831
 QY 315 yValThrArgHisGlnHisGlyArgTyrTrpGlnAlaArgIleGlyArgValaLys 335
 Db 832 TGTTACAAGCATCACCAGCATGGCGGATGGCAAGCAAGGATCGGAAGTGGCTGGCAA 891
 QY 335 nLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAspI 355
 Db 892 CAAGATCTTACCTCGGAACATTCGGCACTGAGAGAGAGCTGCAGAAGCTTATGACAT 951
 QY 355 eAlaAlaLysPheArgGlyLeuThrAlaValThrAspPheAspMetAsnArgTyrAs 375
 Db 952 CGCGCGCATCAAGTTCCTGCGCTGAATGCTGTCCACCACTTGAGATCGCGCGGTATAA 1011
 QY 375 nValLysAlaLysLeuGluSerProSerLeuProIleGlySer---AlaAlaLysArgLe 394
 Db 1012 CGTGAGAGCATATACATGATAGC---AACCTTCCATCGGAGCATGCGCGCAACCGGAG 1068
 QY 394 uLysGluAlaAsn-----ArgProValProSerMetMetMetIle----- 407
 Db 1069 CACCAAGCGCGCTAGAGCTCGCTCGCTCGAGCTCCGCGCAGCCATCGCGCGCACCGA 1128
 QY 408 -SerAsnAsnValSerGluSerGluAsnSerAlaSerGlyTrpGlnAsnAlaAlaValG 427
 Db 1129 AGCGAACACACCGCGCGTGGCGCCACGCTCGACGCTTCCGCTTCCCAATGAA 1188
 QY 427 nHisHisGlnGlyValAspLeuSerLeuHis---GlnHisGlnGluArgTyrAsn-- 445

Db 1189 GTATGATCAGCAGATTACCTGTCTACCTCGCATTCGACGATCACCAGCAAGAACT 1248
 QY 446 -----GlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaAr 460
 Db 1249 TCAGGGCTAGGCTTCGTCTCTACAGCTCAGGTGTAACCTGGAATTCGAAATGCCAA 1308
 QY 460 gala-----CysPhe-----LysGlnGluAs 467
 Db 1309 TGGCAATGGAGCCATGTCAAATTCCTACCAATGCTCATTACATGAGCAGCAGCA 1368
 QY 467 pAspGlnHisPheLeuSerAsnThrGlnSerLeuMetThrAsnIleAspHisGlnSe 487
 Db 1369 ACATCAGCACCAGCATCAGCAGGAGCAGCA-----GATCAACAGGA 1413
 QY 487 rSerValSerAspSerValThrValCysGlyAsnValValGlyTyrGlyGlyTyrG 507
 Db 1414 TGATCAGTCACAAAGCTCCAAACAGCTGCGGTCCATT----- 1453
 QY 507 nGlyPheAlaAlaProVal----- 513
 Db 1454 -CCGTTTGCACACCCATAGCTTTTCAGTGGGTCTATGAAAGCTCCATGACAGCAGCAGG 1512
 QY 514 -----AsnCysAspAlaTyrAlaAlaSerGluPhe-AspTyrAsnA 527
 Db 1513 CACCTTTGGATACCTACCAAAATGTCAGCTTTTCAGACACCAATCTTTGGATGGAAG 1572
 QY 527 laArgAsnHisTyrTyrPheAlaGlnGlnGlnGlnThrGlnGlnSerProGlyGlyAspP 547
 Db 1573 ATTTCGGGAGTGGCACTTTGCAGGCAAAAGGATGAAGAGCCTGATCCTGATCAGAGG 1632
 QY 547 heProAlaAlaMetThrAsnAsnValGlySerAsn 558
 Db 1633 TGATGAGGATCAAGACTGACATAGTTTCTTCTC 1667
 RESULT 15
 AKI01959
 LOCUS
 DEFINITION
 Oryza sativa (japonica cultivar-group) cDNA clone:J033075007, full insert sequence.
 ACCESSION
 AKI01959.1 GI:32987168
 VERSION
 FLI_CDNA; CAP trapper.
 KEYWORDS
 Oryza sativa (japonica cultivar-group)
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 1
 AUTHORS
 The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team., Kikuchi,S., Satoh,K., Negata,I., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Genome Sequencing & Analysis Group., Otonari,Y., Murakami,K., Tida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,I., Kodama,T., Masuda,H., Kobayashi,M., Xie,O., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.
 COLLECTION, mapping, and annotation of over 28,000 cDNA clones from japonica rice
 TITLE
 Science 301 (5631), 376-379 (2003)
 JOURNAL
 MEDLINE
 PUBMED
 2 (bases 1 to 1745)
 REFERENCE
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Ikeda, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itch, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, W., Koya, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE JOURNAL

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, I., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itch, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES
Location/Qualifiers
1. 1745
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clones="J033075007"

ORIGIN

Alignment Scores:

Pred. No.:	6,42e-60	Length:	1745
Score:	963.00	Matches:	233
Percent Similarity:	54.66%	Conservative:	66
Best Local Similarity:	42.60%	Mismatches:	122
Query Match:	30.88%	Indels:	126
DB:	8	Gaps:	19

US-09-980-364-2 (1-579) x AK101959 (1-1745)

Qy	25	SerThrThrThrThrValValValValAlaGlyGlyGlyGlyCysTyrAspProThrAlaAla	44
Db	2	AGCACAACTGCTGCTACATCGAT---CAATAGCTTTTAGTTAT---CTTCCCTCTCA	55
Qy	45	SerAspGluSerSerAlaGlnThrSerPheProSerPheGlyValValValValAsp	64
Db	56	TAAGCTCAACAAACGCTGGTCAAGTCTCTTCAAGATCGGCGCATCATATCTATC	115
Qy	65	AlaPheThrArgAsp---AsnAsnSerHisSerArgAspTyrAspIleAsnGlyCysAla	83
Db	116	GCCATGACATGACATGAGCTGGCTGGCTTATCTCCACATGG---CTCTCTCTCCCTC	172
Qy	84	CysAsnAsnIleHisAsn-----	89
Db	173	TCTAAACAATACCACTGCTGCTCTTGGAGGCGCTCTTACCACATCTGCACTCCACTT	232
Qy	90	-----AspGluGlnAspGlyProLysLeuGluAsnPheLeuGlyArgThrThr	105
Db	233	GGAGAGAGGGGGCCAGCGGCGCTCCGAAGATGGAGGATTTCTC-----	280
Qy	106	ThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySerGlyCysTyrGlyGly	125
Db	281	-----GGCGGC	286
Qy	126	GlyAspGlyGlyGlyGlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGln	145
Db	287	CTAGCGGAGGGCGGCGCGCTCGCGCGCTCCGCGAGCTGCC-----	331
Qy	146	ProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSer	165
Db	332	CGGAGGATCAGCTCAGCTGCGCGGAGCTGTGTAGCATCGCGCGGGTTC-----	382
Qy	166	MetAsnSerSerThrSerCysAspAsnAsnAspSerAsnAsnValAlaGln	185
Db	383	TTGCGCGGTATCCCGCGCTCGAGAAC-----GCCGCG	415
Qy	186	GlyLysThrIleAspAspSerValGluAlaThr-----ProLysLys	199
Db	416	GGGTGACCATCGCAATGGCGACCGCGGGGAGCTGGCGATCCCGCGGAGG	475
Qy	200	ThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTrp	219
Db	476	ACCGCGAGACGCTTCGGGACACGCGACGCTTACCGTGTGTGCTAGCAGCGGTGG	535
Qy	220	ThrGlyArgTyrGluAlaHisLeuTyrAspAsnSerCysLysArgGluGlyGlnThrArg	239
Db	536	ACGGGGAGGTAGAGCGCGACCTGTGGACAAATAGTCCCGCGGAGGCGAAGCGCGC	595
Qy	240	LysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyr	259
Db	596	AAAGCGCGCAAGTCTACTTAGGAGGTTATGATAGGAGGAGAGGCGCGCAAGAGCTTAC	655
Qy	260	AspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrThrThrThrThrThrThr	279
Db	656	GGCTTCGCGCGCTTAAAGTCTGGGGTCCACACACAGCAAACTTCCCTGTGTGCCAAC	715
Qy	280	TyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeu	299
Db	716	TACGAGACGGAGCTGGAGAGATGAATGCCATGACCGCGGAGGAGTTTCATCGCTGCTG	775
Qy	300	ArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHis	319
Db	776	CGCAGAAAGACGCGCGCTTCTCAAGAGGGGCTTCCATCTACAGAGGAGTAAACAGACAT	835
Qy	320	HisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyr	339
Db	836	CATCAGCAGCGCGGTGGCAAGCGAGGATCGCAGGTTGCGCGAACAAGACCTGTAC	895
Qy	340	LeuGlyThrPheGlyThrGlnGluAlaGluAlaGluAlaTyrAspIleAlaAlaIleLys	359
Db	896	TTGGACATTTTCAGCAGCAGGAGGCTGCCGAGGCTACGACATTCGCGCTATCAAG	955
Qy	360	PheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIle	379

[illegible]

Search completed: July 19, 2004, 17:19:55
Job time : 6646 secs

PT Brassica napus BNM3 DNA sequences, useful for transforming host cells to
 PT produce asexually derived embryos.

XX Claim 4; Fig 2; Sipp; English.

XX The present sequence is given in a specification relating to a Brassica
 CC napus microspore embryo (BME) gene obtained during the induction of
 CC microspore embryogenesis. The isolated DNA molecule can be used in a
 CC vector that is used to transform plant cells. The vector can be used in
 CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant

XX SQ Sequence 2014 BP; 596 A; 411 C; 462 G; 545 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,28e-268 Length: 2014
 Score: 3119.00 Matches: 579
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-980-364-2 (1-579) x AAC91399 (1-2014)

QY 1 MetAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
 DB 120 ATGAATAATACTGGTTAGGCTTTTCTCTCTCTATGACAAATAATCACCATCGTAAG 179
 QY 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
 DB 180 GACGCTCTACTCTTCACCAACACACACACCTGCTAGATGTCGCCGAGAGTACTGTTACGAT 239
 QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
 DB 240 CCGACCGCTGCTCCGATGAGTCTTCAGCCATCAACATCGTTTCTCTCCCTTTGGT 299
 QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsn 80
 DB 300 GTGCTGCTGATGCTTTTACCAGAGACAAACATAGTCACTCCCGAGATTGGGACATCAAT 359
 QY 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPhe 100
 DB 360 GGTTCGTGATGCAATACATCCACACGATGAGCAAGATGGACCAAGCTTGAGAAATTC 419
 QY 101 LeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySer 120
 DB 420 CTGTCGCCGACACACACGATTTACACACCAACGAAACGTTGGAGATGGAAGTGAAGT 479
 QY 121 GlyCysTyrGlyGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLysThr 140
 DB 480 GGTGTTATGAGGAGGAGACGGTGTGTGGTCTCACTAGGACTTTCGATGATAAGACA 539
 QY 141 TrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys 160
 DB 540 TGGCTGAGAAATCAACCGCTGATATGTTGATATCAAGAAATGCGCAATGCTGCAAAA 599
 QY 161 GlyLeuSerLeuSerMetAsnSerThrSerCysAspAsnAsnAsnAspSerAsnAsn 180
 DB 600 GGCCTGTCCTCTCAATGAATCACTACTCTTCTGTGATACAAACAACGACGACCAATAC 659
 QY 181 AsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrProLysLysThr 200
 DB 660 AACGTTGTTGCCCCAAGGAGGAGACTATTGATGATAGCGTTCAAGCTACACCGAAGAAACT 719
 QY 201 IleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTyrThr 220
 DB 720 ATTGAGAGTTTGGACAGAGAGCTCTATATACCGCGTGTTCACAGGATCCGTTGGACA 779
 QY 221 GlyArgTyrGluAlaHisLeuTyrAspAsnSerCysLysArgGluGlyGlnThrArgLys 240
 DB 780 GGAAGATATGAGGCACATTTATGGATAATAGTTGTAAAGAGAGAGGCCCAACCGCCAAA 839
 QY 241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAsp 260

DB 840 GGAAGCAAGTTTATTGGGAGGTTTATGACAAAGAAAAAGACAGCTAGGGCTTATGAT 899
 QY 261 LeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyr 280
 DB 900 TTAGCCGCACTCAAGTATTGGGGAACCACTACTACTACTTCCCATGAGCGAATAT 959
 QY 281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300
 DB 960 GAAAGAGAGTGAAGAGATGAAGCATGACAAAGCAAGAGTATGTTGCCCTCACTGCGC 1019
 QY 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320
 DB 1020 AGGAAAGTAGTGGTTTCTCTCGTGTGCTCGATTTATCTGGAGTAAACAGACATCAC 1079
 QY 321 GlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 340
 DB 1080 CAACATGGAAGATGGCAAGCTAGGATAGGAGTCCCGGTAAACAAAGACCTCTACTTG 1139
 QY 341 GlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAspIleAlaAlaLysPhe 360
 DB 1140 GGAACCTTTTGGCACACAAAGAGAGCTGCGAGGAGTATCGACATTCGGGCCATCAATTC 1199
 QY 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeu 380
 DB 1200 AGAGGATTAACCGCAGTGAATTCGACATGAACAGATACACAGTTAAAGCAATCCTC 1259
 QY 381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro 400
 DB 1260 GAAAGCCCTAGTCTTCTTATGTTAGCGCCGCAAAACGTTCTCAAGAGGCTTAACCGTCCG 1319
 QY 401 ValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGly 420
 DB 1320 GTTCCAAGTATGATGATGATCACTAGTAAATACGTTTCAGAGAGTGAAGTAGTCTAGCGGT 1379
 QY 421 TrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeuHisGlnHis 440
 DB 1380 TGGCAAAACGCTGCGGCTTCAGCATCATCAGGAGTAGATTTTGAAGTTATTGCACCAACAT 1439
 QY 441 GlnGluArgTyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArg 460
 DB 1440 CAAGAGAGTACAAATGGTTTATTATTAATGAGAGAAACTTCTCTTCGAGAGTGTAGG 1499
 QY 461 AlaCysPheLysGlnGluAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMet 480
 DB 1500 GCTTGTGTTCAACCAAGAGAGTATCAACACCATTTCTTTGAGCAACACGACGAGGCTCATG 1559
 QY 481 ThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnVal 500
 DB 1560 ACTAATATCGATCATCAAGTCTGTTTCGGATGATTCGGTTACTGTTTGGGAAATGTT 1619
 QY 501 ValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAlaAla 520
 DB 1620 GTTGGTTATGGTGTATCAAGGATTTTCAGCCCGGTTAACTGCGATGCTCGCTCGCTGCT 1679
 QY 521 SerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
 DB 1680 AGTGAGTTTGATTATACGCAAGAAACCATATTACTTTGCTCAGCAGCAGACGCCAG 1739
 QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
 DB 1740 CAGTCGCCAGGTGGAGATTTTCCCGCGCAATGACAAATATGTGGCTCTTAATATGAT 1799
 QY 561 TyrHisGlyGluGlyGlyGlyGluValAlaProThrPheThrValTyrAsnAsnAsn 579
 DB 1800 TACCATGGGAGGTTGGTGGAGAGTTGCTCCAACATTTACAGTTTGGACGCAAT 1856

RESULT 2

AAC91399 standard; cDNA; 2011 BP.

ID AAC91399

XX AC

XX AAC91399;

XX AC

DT 19-MAR-2001 (first entry)
 XX Brassica napus BNM3B cDNA.
 XX
 KW Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3B;
 KW asexually derived embryo production; increased regenerative capacity; ss.
 XX Brassica napus.
 OS
 XX EP1057891-A1.
 PN
 XX 06-DEC-2000.
 PD
 XX 02-JUN-1999; 99EP-00201745.
 PF
 XX 02-JUN-1999; 99EP-00201745.
 PR
 XX (CPRO-DLO CENT PLANTENVERDEDELINGS REPROD.
 PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
 PA
 XX Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
 PI Boutillier K;
 PI
 XX WPI; 2001-026720/04.
 DR
 DR P-PSDB; AAB50873.
 XX
 PT Brassica napus BNM3 DNA sequences, useful for transforming host cells to
 PT produce asexually derived embryos.
 XX
 XX Claim 4; Fig 2; 51pp; English.
 PS
 CC The present sequence is given in a specification relating to a Brassica
 CC napus microspore embryo (BNM) gene obtained during the induction of
 CC microspore embryogenesis. The isolated DNA molecule can be used in a
 CC vector that is used to transform plant cells. The vector can be used in
 CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant
 CC
 XX Sequence 2011 BP; 610 A; 395 C; 462 G; 544 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 3,67e-260 Length: 2011
 Score: 3031.00 Matches: 564
 Percent Similarity: 98.27% Conservative: 5
 Best Local Similarity: 97.41% Mismatches: 10
 Query Match: 97.18% Indels: 0
 DB: 4 Gaps: 0

US-09-980-364-2 (1-579) x AAC91399 (1-2011)
 Qy 1 MetAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluInAsnHisHisArgLys 20
 Db 51 ATGAATAAATCTGTTAGGCTTTCTCTCTCTCTATGAACAAATCACCATCGTAAG 110
 Qy 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
 Db 111 GACGCTCTGCTCTCCACACACACACCGCGGTAGATGTCGCGGAGAGTACTCTTAGAT 170
 Qy 41 ProThrAlaAlaSerAspGluSerSerAlaAlaGlnThrSerPheProSerProPheGly 60
 Db 171 CCGACCGCTGCTCGATGAGCTCTCAGCATCCAAACATCGTTCTCTCTCTCTCTTGGT 230
 Qy 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsn 80
 Db 231 GTCGCTCTCGATGCTTTCCACAGACACAAATAGTCACTCCCGAGATTGGGACATCAAT 290
 Qy 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluInAspGlyProLysLeuGluAsnPhe 100
 Db 291 GGTAGTGCATGTATTAACATCCAAATGATGAGCAAGATGAGCACCACAAACTTGAGAAATTC 350
 Qy 101 LeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySer 120
 Db 351 CTGTGCCGACACACACGATTTACAAACACCAACCAACGTTGGAGATATCGATGGAAGT 410

121 GlyCysTyrGlyGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLysThr 140
 411 GCGTGTATGGAGGAGGAGCGGTGGTGGCTCACTAGGACTTTCGATGATAAAGACA 470
 141 TrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys 160
 471 TGGCTGAGAANTCAACCCGCGGATAATGTTGATAATCAAGAAATGCAATGGTGCAAAA 530
 161 GlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsnAsn 180
 531 GGCCTGTCCCTCTCAATGAACCTCATCTTCTGTGATAACAACAACACTACAGCAGTAAC 590
 181 AsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrProLysLysThr 200
 591 AACCTTGTGCCCAAGGAGAGACTTATGATGATAGCGTTGAGGTACACCGAAGAACT 650
 201 IleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTrpThr 220
 651 ATTGAGAGTATTGGACAGAGGACGCTATATACCGCGGTGTTCAAGGCATCGGTGGACA 710
 221 GlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLys 240
 711 GGAGATATGAGGCACATTTATGGGATATAGTTGTAACGAGAGGCGCAACCGCGCAA 770
 241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAsp 260
 771 GGAACACAAGTTTATTGGGAGGTTATGACAAAGAGAAAAGCAGCTAGGCTTATCAT 830
 261 LeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyr 280
 831 TTAGCCGCACTCAAGTATTGGGAAACCCACTACTACTACTTCCCCCATGAGCGAATAT 890
 281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300
 891 GAGAAAGAGATAGAGAGATGAGGCACATGACAGGCAAGAGTAGTTGCTCTCACTTCG 950
 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320
 951 AGCAAAAGTAGTGGTTTCTCTCGTGCATCGATTTATCGTGGAGTAAACAGACATCAC 1010
 321 GlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 340
 1011 CAACATGGAAGATGCAAGCTAGGATAGAGAGTCGCGGTAAACAAGACCTCTACTTGT 1070
 341 GlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaLysPhe 360
 1071 GGAACCTTTGGCACACAAGAAAGCTGCAGAGGCATACGACATTCGCGCCATCAAAATTC 1130
 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaLysLeu 380
 1131 AGAGGATTAACCGCAGTGAACCTTTCGACATGAACAGATACACGTTAAAGCAATCCCTC 1190
 381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro 400
 1191 GAAAGCCCTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1250
 401 ValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGly 420
 1251 GTTCCAAGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1310
 421 TrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeuHisGlnHis 440
 1311 TGGCAAAACCTTCGCGTTCAGCATCATCAGGAGTAGATTGAGCTTATTGCGAGCAACAT 1370
 441 GlnGluArgTyrAsnGlyTyrTyrTrpAsnGlyGlyAsnLeuSerSerGluSerAlaArg 460
 1371 CAAGAGAGTACAATGGTTTATTATTACATGAGGAAATCTGCTCTTCGAGAGTGTCTAGG 1430
 461 AlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMet 480
 1431 GCTTGTTCACAAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1490

QY 481 ThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnVal 500
 Db 1491 ACTAATATCGATCATCAAAAGTTCTGTTTCAGATGATTCGGTTACTGTTTGTGGAATGTT 1550
 QY 501 ValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAlaAla 520
 Db 1551 GTTGTTTATGGTGTATCAGGATTCGACCCCGGTTAACTGCGATGCCCTACGCTGCT 1610
 QY 521 SerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
 Db 1611 AGTGAGTTTGACTATAACGCAAGAAACCAATTACTTGTCTCAGCAGCAGCAGCCAG 1670
 QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
 Db 1671 CATTCGCCAGGAGGAGATTTTCCCGCGCAATGACGAATAATGTTGGCTCTAATATGTA 1730
 QY 561 TyrHisGlyGlyGlyGlyGlyGluValAlaProThrPheThrValTyrAsnAspAsn 579
 Db 1731 TACCATGGGAAGTGTGTGGAGAGTTGCTCCAACTTACAGTTTGGAAACGCAAT 1787

RESULT 3
 AAC91400
 ID AAC91400 standard; DNA; 4873 BP.
 AC
 XX AAC91400;
 DT 19-MAR-2001 (first entry)
 XX
 DE Brassica napus BMN3A DNA sequence.
 XX
 KW Brassica napus microspore embryo; BMN; microspore embryogenesis; BMN3A;
 KW asexually derived embryo production; increased regenerative capacity; ds.
 OS Brassica napus.
 XX
 PN EP1057891-A1.
 XX
 PD 06-DEC-2000.
 XX
 PF 02-JUN-1999; 99EP-00201745.
 XX
 PR 02-JUN-1999; 99EP-00201745.
 XX
 PA (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
 PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
 PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
 PI Boutilier K;
 XX
 DR WPI; 2001-026720/04.
 DR P-PSDB; AAB50872.
 XX
 PT Brassica napus BMN3 DNA sequences, useful for transforming host cells to
 PT produce asexually derived embryos.
 XX
 PS Claim 1; Page 32-33; 51pp; English.
 XX
 CC The present sequence is given in a specification relating to a Brassica
 CC napus microspore embryo (BMN) gene obtained during the induction of
 CC microspore embryogenesis. The isolated DNA molecule can be used in a
 CC vector that is used to transform plant cells. The vector can be used in
 CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant
 XX
 SQ Sequence 4873 BP; 1596 A; 856 C; 907 G; 1514 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,6e-229 Length: 4873
 Score: 2690,50 Matches: 574
 Percent Similarity: 53.24% Conservative: 1
 Best Local Similarity: 53.15% Mismatches: 4
 Query Match: 86.26% Indels: 505
 Gaps: 4

US-09-980-364-2 (1-579) x AAC91400 (1-4873)
 QY 1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
 Db 1620 ATGAATTAATACCTGGTTAGGCTTTCTCTCTCTCTATGAACAAATCACCATCGTAAG 1679
 QY 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGlyTyrCysTyrAsp 40
 Db 1680 GACGTCTACTCTTCCACCACCAACCGTCGTAGATGTCGCGGAGAGTACTGTACCAT 1739
 QY 41 ProThrAlaAlaSerAsnGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
 Db 1740 CGACCGCTGCTCCGATGAGTCCTCAGCCATCCAACATCGTTTCTCTCCCTTTGGT 1799
 QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArg----- 75
 Db 1800 GTCGTCGTGATGCTTTCACCAGACACAACTAGTCACTCCCGAGGTTATTGTTTAGA 1859
 QY 75 ----- 75
 Db 1860 ACTACTTGTTTTTTTGTGATTGTTTATTGTTAGTTTCTCTCTTCTTCCAAATCGGTAGA 1919
 QY 75 ----- 75
 Db 1920 ACAAGACCAATACACACGCGCATAGCCCTATTTTCTCTGGGCTTATTATCG 1979
 QY 75 ----- 75
 Db 1980 ATTTCATTTTATTGAGAATATCAATGTGCGGGTTGATGTTGTTGTCATATAGTAAT 2039
 QY 75 ----- 75
 Db 2040 ACTAAACATATGCCAGTTTATACATAGATTTTTTTAAAGATATACATGGATGAATG 2099
 QY 75 ----- 75
 Db 2100 AAATTTGACATTTCTCTCTTTTATTCAATATCATATATGATCACATACATGTACCTTT 2159
 QY 75 ----- 75
 Db 2160 TGATTTGTATATTGTTTCTTACATGTTGAAGAGAGAAATACCAAAATACCCATTTGTATA 2219
 QY 75 ----- 75
 Db 2220 TTATAGTCGGTGATCAAAAGTAAATTTAAACAAATTATGATAATATAGGCCATTAATCTT 2279
 QY 76 -----AspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsn 89
 Db 2280 TGATTTTCTTTTATAGATTGGGACATCAATGTTGTGATGCAATAACATCCACAC 2339
 QY 90 AspGluGlnAspGlyProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsn 109
 Db 2340 GATGAGCAAGATGGACCAAGCTTGAGAAATTTCTTGGCCGACACACCATTTACAAC 2399
 QY 110 ThrAsnGluAsnValGlyAspGlySerGlyCysTyrGlyGlyGlyAspGlyGly 129
 Db 2400 ACCAACGAAAACGTTGGAGATGGAAAGTGGCTGTTATGGAGGAGGAGCGGTGTT 2459
 QY 130 GlyGlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsn 149
 Db 2460 GGTGGCTCACTAGGACTTTCGATGATAAAGACATGGCTGAGAAATCAACCGTGGATAAT 2519
 QY 150 ValAspAsnGlnGluAsnGlyAsnAlaLysGlyLeuSerLeuSerMetAsnSerSer 169
 Db 2520 GTTGATAATCAAGAAATGGCAATGCTGCAAAAGGCTGTCCCTCTCAATGAATCATCT 2579
 QY 170 ThrSerCysAspAsnAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThrIle 189
 Db 2580 ACTTCTTGTGATACAAACAGACAGCAATTAACACGTTGTGTGCCAAGGAGAACATATT 2639
 QY 190 AspAspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSer 209

RESULT 4					
ID	ACC00801				
XX	ACC00801 standard; cDNA; 2598 BP.				
XX	AC				
XX	ACC00801;				
DT	16-MAY-2003 (first entry)				
XX	Glycine max oil trait related cDNA sequence SEQ ID NO:351.				
DE	XX				
XX	Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase; receptor-like protein kinase; mitogen activated protein kinase; oil; KW				
KW	LIP15-like transcription factor caleosin; ATP citrate lyase; SNF1; KW				
KW	CYC-like transcription factor; antisense inhibition; co-suppression; KW				
KW	transgenic plant; gene; ss.				
XX	OS				
OS	Glycine max.				
XX	WO2003002751-A2.				
PX	09-JAN-2003.				
XX	27-JUN-2002; 2002WO-US020152.				
PF	29-JUN-2001; 2001US-0301913P.				
XX	(DUPO) DU PONT DE NEWMOURS & CO E I.				
PA	(PION-) PIONEER HI-BRED INT INC.				
PA	Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT; Jones TJ, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H, Shen B; Tarczynski MC;				
PI	WPI; 2003-201509/19.				
XX	P-FSDB; ABR40766.				
DR	Novel nucleotide fragment encoding polypeptides having receptor-like protein kinase activity, caleosin-like activity, useful for altering oil phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.				
PT	Claim 18; Page 368; 542pp; English.				
PS	The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (PP) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, Lip15-like transcription factor activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like activity, and CYC-like transcription factor activity. Also described: (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00801 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention				
XX	Sequence 2588 BP; 863 A; 513 C; 554 G; 658 T; 0 U; 0 Other;				
SQ	Alignment Scores:				
	Prod. No.: 1.19e-106 Length: 2588				
	Score: 1310.00 Matches: 322				
	Percent Similarity: 56.52% Conservative: 77				
	Best Local Similarity: 45.61% Mismatches: 139				
	Query Match: 42.00% Indels: 168				
	DB: 7 Gaps: 29				
US-09-980-364-2 (1-579) x ACC00801 (1-2588)					
YV	4 AsnTrpLeuGlyPheSerLeuSerProTyrGluInAsnHisArgLysAspValTyr 23				

1092	TTCGAGGAAGAGTAGTGGGTTTTTCGCGGTGCATCCATTTATCGAGGAGTGCACGAGA	1155
319	HiHisGlnHisGlyArgTTPGlnAlaAraGileGlyArgValaIaGlyAsnLysAspLeu	338
1152	CACCACCAACATGGAAGGTGGCAAGGAGGAGATTGGAAGAGCTTGCTGGCAACAAGGATCTT	1211
339	TyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAspIleAlaAlaIle	358
1212	TACTTGGGAACCTTTTAGCACCCCAAGAGAGCGCAGCGAAGCATATGATGTAGCAGCAATC	1271
359	LysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAla	378
1272	AAATCCGAGGACTAAGTGTGTTACAAACTTTGACATGACGAGATATGACGTGGAAGC	1331
379	IleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGlu-----	396
1332	ATACTTTGAGAGCACCACTTTGCCAATAGGTGGTCTGCAAAAGCGTTTGAAGAGTATGGAG	1391
397	-----AlaAsnArgProValProSerMetMet	405
1392	CAGTTTCAACTGAGTGGGATAATGGTCATAGAGCAGATCAAGTAGATCATATGATCATC	1451
406	MetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGlyTrpGlnAsnAlaAla	425
1452	ATGAGTCTCTCACCTAATCAAGGAATCAATAACAACCTATGACGAGGCGGGAACAGCAACT	1511
426	ValGlnHisHisGlnGlyValAspLeuSerLeuLeuHisGlnHisGlnGluAraGlyTyrAsn	445
1512	-----CATCATAACTGGCAATGCTCATGCTATGCCAACCTCAACTTGGCACCAACC	1565
446	GlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaAlaGlyPheLysGln	465
1566	ATGCATACCTTATGACAAAGAAATTAAT-----TGGTGCAAGCAAGAA	1610
466	GluAspAspGln-----	489
1611	CAACAAGACACACTGTGATGCCCTCACTCTTTGCTTATTCAGATATTCATCAACTTCAG	1670
470	-----HisHisPheLeuSerAsnThrGlnSerLeu-----MetThr	481
1671	CTAGGGAAACAATGGAAACACATACTCTTTCCACAAATTCAGGGTTGCACCCTATGTTG	1730
482	AsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnValVal	501
1731	AGCATGGAT-----TCTGCTTCCATTCAGAAATAGCTCTCTTCTTAACCTCGGTTGTT	1781
502	-----GlyTyrGlyGlyTyrGlnGlyPhe-----AlaAlaProVal---AsnCysAspAla	517
1782	TATGATGGTTATGAGGTGGTGGGGGCTACAACTGTGATGCTATGTGGAACTACTACTGCT	1841
518	TyrAlaAlaSerGluPheAspTyrAsn-----	526
1842	GTGTGTCGAAGTGATGTGTATCAAATCCCAAGAACCAATCATGTTTTGGTGATAATGAG	1901
526	-----	526
1902	ATAAAGCACTGGTTATGAAGAGTGTATGGCTCTCGCAACTGATCTTTATCATGCAACAT	1961
527	AlaArgAsnHisTyrTyrPheAlaGlnGlnGlnGlnThrGlnGlnSerProGlyGlyAsp	546
1962	GCAAGGAACTGTATTATCTTACTCAACAGCAA-----	1994
547	PheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGly	566
1995	-----TCATCTCTGTTGATACAGCTGAAGGCTAGTCATATGATCAAGGCTCTGCATGC	2048
567	GlyGluValAlaProThr	572
2049	AATACTGGGTTCCAACT	2066

RESULT 5
ACC00803
ID ACC00803 standard; cDNA: 2463 BP.

QY 20 -----LysAspValTyrSerSerThrThrThrVal 30
Db 257 ACGGACACCTTCTCGTTTGGTTTCACCTGATGAATCTCAAGCACT----- 304
QY 31 ValAspValAlaGlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerAla 50
Db 305 ---GATGTAGCAGGAGAC---TGCTTTGATCTCACTTCT-----GACTCAACTCCT 349
QY 51 IleGlnThrSerPheProSerPheGlyValValValAspAlaPheThrArgAspAsn 70
Db 350 CATTTACTCAACCTTCCCTCT---TACGCG---ATATACGAAGCTTTTCATAGAGCAAC 403
QY 71 AsnSerHis---SerArgAspTyrAspIleAsn-----Gly 81
Db 404 AATATTCAACCACTCAAGATTGAAGGAGAACTACAACACCCAAAACCTTGCTATTGGGA 463
QY 82 CysAlaCysAsnAsn-----IleHisAsn-----AspGluGlnAspGlyPro 95
Db 464 ACTTCATGCAGCAACCAAAACATGAACCAACCAACCATCAGCAACCAACCAACACAGCCA 523
QY 96 LysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsn---Val 114
Db 524 AAGCTTGAATACTTCTCGGTGACACTCA-----TTTGGTGAACATGAGCAACCCCTAC 577
QY 115 GlyAspGlySerGlySerGlyCysTyr-----GlyGly 125
Db 578 GGTGTAACCTCAGCCTCTACAGAATACTATGTTCCCGGCTACGCCGGTATTGGCCGGTGGC 637
QY 126 GlyAspGlyGly-----GlyGlySerLeuGlyLeuSerMet 137
Db 638 GCGCGCGTGTAGCAATAGCAGCAACACACAGCACTAGTCTCATAGGGTATCCATG 697
QY 138 IleIysThrTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnIleAsnGlyAsn 157
Db 698 ATAAGACATGTTGAGGAACCAACCAACCACTCAGAAAACCAATAACAACAACAAAT 757
QY 158 AlaAlaLysGly-----LeuSerLeuSerMetAsn--- 167
Db 758 GAAGTGTGTGCAATAGTAGAAGCAGTGTGCAGCAGACTCATCACTTTTCATAGGACT 817
QY 168 -----SerSerThr-----SerCysAspAsnAsnAsp 177
Db 818 GGTTCACAATCAAGCACATCACTACCCCTTCTCACTGCTAGTGTGGATAATAGAGAGAGT 877
QY 178 SerAsnAsnAsn-----ValValalaGlnGlyIysThrIleAspAspSer 192
Db 878 TCTTCTGATAACAACCAACCACTACCGCTGCCTTGATACAACCCCAACCGGAGCC 937
QY 193 ValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArg 212
Db 938 ATTGAAACTGCACCAGAAAGTCCATTGACACTTTTGGACAGAGAACTTCTATCTACCGT 997
QY 213 GlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCys 232
Db 998 GGTGTAAACAGGCATAGGTGACCGGAGGTATGAGCTCACCTGTCGGGATAATAGTTGT 1057
QY 233 LysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGlu 252
Db 1058 AGAAGAGAGGACAACTCGCAAGGAGGCAAGTTCATTGGGAGGTATGACAAGAA 1117
QY 253 GluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyrTrpGlyThrThrThr 272
Db 1118 GAAAGGCGAGCTAGAGCCTAGATTGGCAGCACTAAATACTGGGNAACAACCTACGACA 1177
QY 273 ThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLysHisMetThrArg 292
Db 1178 ACAATTTTCCAAATAGCCACTATGAGAGAGAGTGTGAAGAAATGAAGACATGACTAGG 1237
QY 293 GlnGluTyrValalaSerLeuArgLysSerSerGlyPheSerArgGlyAlaSerile 312
Db 1238 CAAGAGTACGTTGGCTCATTCAGGAAGGAGAGTAGTGGGTTTTCTCGCGGGCATTCATT 1297
QY 313 TyrArgGlyValThrArgHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgVal 332

Db 1298 TATCGAGGTGTCGAGAGACACCATCAACATGGAAGATGGCAAGCGAGGATTGGAAGATT 1357
QY 333 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaGluAla 352
Db 1358 GCTGGCAACAAGGATCTCTACTTGGAACTTTTCAGCACCCCAAGAGAGCGCAGCAGAACA 1417
QY 353 TyrAspIleAlaAlaLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsn 372
Db 1418 TATGATGTAGCAATCAAAATTCAGAGCACTAGTGTGTGTACAACTTTTGACATGAGC 1477
QY 373 ArgTyrAsnValLysAlaLysLeuGluSerProSerLeuProIleGlySerAlaAlaLys 392
Db 1478 AGATATCACTGAAAAGCATACTTGAGAGCACCACTTTGCCAAATTTGGTGTGCAAG 1537
QY 393 ArgLeuLysGlu-----AlaAsnArg 399
Db 1538 CTTTTGAAGGATATGAGCAGCGTGAAGCTGAGGTGGAGAATGTTCTAGAGCAGATCAA 1597
QY 400 ProValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSer 419
Db 1598 GAAGATCATAGTAGCATCATGAATCTCACTTAACCTCAAGGAATCATTAACAACATGCA 1657
QY 420 GlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGln 439
Db 1658 GCAGGAGGAACAACAGCGCACTCATCATCACTAAGTGCACAATGCTCTTGCAATTCACCAA 1717
QY 440 HisGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAla 459
Db 1718 CCTCAACCTTGCAACCACTACCTTATGACAAAGAAATTAAT----- 1765
QY 460 ArgAlaCysPheLysGlnGluAspAspGln----- 469
Db 1766 -----TGTGCAAGCAAGAAACAAGCAACTCTGATGCTCTCACTCTTTGCTATTATCA 1819
QY 470 -----HisHisPheLeuSerAsnThrGln 477
Db 1820 GATATTCACTACAGTAGGGAACAAATGGCACACAACTTTTTCACAAATCA 1879
QY 478 SerLeu-----MetThrAsnIleAspHisGlnSerSerValSerAspAspSerValThr 495
Db 1880 GGGTGTGACCCCTATGTTAAGCATGAT-----TCTGCTTCCATTGACAATAGCTCT 1930
QY 496 ValCysGlyAsnValVal-----GlyTyrGlyGlyTyrGlnGlyPhe---AlaAlaPro 512
Db 1931 TCATCTACTCTGTTGTTTATGATGTTATGAGGTGTTGGGGCTATAATGTGATTCCT 1990
QY 513 Val---AsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsn----- 526
Db 1991 ATGGGGACTACTACTACTGTTGTTGCAAAATGATGTGATCAAAATCCAAAGCAATCAT 2050
QY 526 ----- 526
Db 2051 GGTTTTGGTGATAATCAGATAAGCACTTGTTATGAAGTGTGTATGTTCTTACAACT 2110
QY 527 -----AlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
Db 2111 GATCCTTATCATGCACATGCAAGGAACCTGTTATTATCTTACTCAACAGCAACCATCTCT 2170
QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
Db 2171 -----GTTGATGAGTGAAGGCTAGTGCATAT 2197
QY 561 TyrHisGlyGluGlyGlyGluValAlaProThr 572
Db 2198 GATCAAGGATCTGCATGCAATACTTGGGTTCCAACT 2233
RESULT 6
ACCC0806
ID ACC00806 standard; cdna; 1990 BP.
XX
AC ACC00806;
XX

16-MAY-2003 (first entry)

Glycine max oil trait related cDNA sequence SEQ ID NO:361.

Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase; receptor-like protein kinase; mitogen activated protein kinase; oil; lipid-like transcription factor caleosin; ATP citrate lyase; SNF1; CKC-like transcription factor; antisense inhibition; co-suppression; transgenic plant; gene; ss.

Glycine max.

WO2003002751-A2.

09-JAN-2003.

27-JUN-2002; 2002WO-US020152.

29-JUN-2001; 2001US-0301913P.

(DUPO) DU PONT DE NEMOURS & CO E. I. (PION-) PIONEER HI-BRED INT INC.

Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT, Jones TJ, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H, Shen B, Tarczynski MC,

WPI; 2003-201509/19.

P-PSDB; ABR40771.

Novel nucleotide fragment encoding polypeptides having receptor-like protein kinase activity, caleosin-like activity, useful for altering oil phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.

Claim 18; Page 381-382; 542pp; English.

The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (PP) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, lipid-like transcription factor activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like activity and CKC-like transcription factor activity. Also described: (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a plant such as corn, soybean, wheat, rice, altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00636 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention

Sequence 1990 BP; 628 A; 391 C; 459 G; 512 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8e-90	Length:	1990
Score:	1119.50	Matches:	275
Percent Similarity:	55.70%	Conservative:	77
Best Local Similarity:	43.51%	Mismatches:	147
Query Match:	35.89%	Indels:	133
DB:	7	Gaps:	23

US-09-980-364-2 (1-579) x ACC00806 (1-1990)

1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20

126 ATGACACACACACTGGCTTTCGTTCCCTCTTCT----- 158

21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40

158 ----- 158


```
Db 366 ATGAGCACATGACTAGGCAAGTAGTACGTTGGCTCATTGAGAGGAGAGTAGTGGTTT 425
Qy SerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGln 326
Db 426 TCTCGGGGTGCATCCATTTATCGAGAGGTGACGAGACACCAACATCGAAGGTGGCAA 485
Qy 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
Db 486 CGAGAGATGGAAGAGTTGCTGCAACAGGATCTTTACTTGGAACTTTTAGCACCCAA 545
Qy 347 GluGluAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
Db 546 GAAGAGCGAGCGAAGCATATGATGATAGCAGCAATCAAAATTCGAGGAGTAAGTGTGTT 605
Qy 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
Db 606 ACAACTTTGACATGACGATATGCTGGAAGCATATCTTGAGAGCACCACCTTTGCCA 665
Qy 387 IleGlySerAlaAlaLysArgLeuLysGlu----- 396
Db 666 ATAGTGGTCTGCAAGCGCTTTGAAGGATATGGAGCAGGTTGAACTGAGTGTGGATAAT 725
Qy 397 -----AlaAsnArgProValProSerMetMetIleSerAsnAsnValSerGlu 413
Db 726 GGTATAGACGACATCAAGTAGATCATGATATCATCATGAGTTCTCACCTAACTCAAGGA 785
Qy 414 SerGluAsnSerAlaSerGlyTrpGlnAsnAlaValGlnHisGlnGlyValAsp 433
Db 786 ATCAATACAACTATGACAGAGGGGACAGCACT-----CATCATTAACGGCACAT 839
Qy 434 LeuSerLeuLeuHisGlnHisGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsn 453
Db 840 GCTATGCAATTCACCAACCTCAACCTTGACACCACTGACACCTATGATGACAAAGA 899
Qy 454 LeuSerSerGluSerAlaArgAlaCysPheLysGlnGluAspAspGln----- 469
Db 900 ATTAAT-----TGGTGCAAGCAAGCAACAGCAACTCTGTATGCCCT 944
Qy 470 -----HisHis 471
Db 945 CACTCTTTGTATTATTCAGATATTCATCACTTCAGCTAGGGAACAATGGAACATAAC 1004
Qy 472 PheLeuSerAsnThrGlnSerLeu-----MetThrAsnIleAspHisGlnSerSerVal 489
Db 1005 TTCTTTCACCAAAATTCAGGGTGCACCCCTATGTGAGCATGGAT-----TCTGCT 1055
Qy 490 SerAspAspSerValThrValCysGlyAsnValVal-----GlyTyrGlyGlyTyrGln 507
Db 1056 TCCATTGACATAGCTCTCTCTTAACCTCGGTGTTTATGATGCTTATGAGGTGGTGG 1115
Qy 508 GlyPhe---AlaAlaProVal---AsnCysAspAlaTyrAlaAlaSerGluPheAspTyr 525
Db 1116 GGCTACAATGTGATGCCTATGGGAACCTACTACTGCTGTTGTTGCAAGTGTGTGATCAA 1175
Qy 526 Asn----- 526
Db 1176 ATCCAAAGAACATCATGTTTGGTGATATGAGTAAGAACACTTGGTTATGAAGT 1235
Qy 527 -----AlaArgAsnHisTyrTyrPheAla 534
Db 1236 GTGTATGGCTGTGCAACTGATTTCTATCATGCATGCAAGGAACTTGTATTATCTTACT 1295
Qy 535 GlnGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsn 554
Db 1296 CAACGCAA-----TCATCTCTGTGTATACA 1322
Qy 555 ValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaProThr 572
Db 1323 GTGAAGGCTAGTGCATATGATCAAGGGTCTGCATGCAATACTTGGGTTCCAACT 1376
RESULT 8
ADD30878
ID ADD30878 standard; cDNA; 1818 BP.
```

```
XX AC ADD30878;
XX DT 15-JAN-2004 (first entry)
XX DE Plant yield-related polynucleotide clone G1793.
XX KW ds; transcription factor; transgenic plant; growth rate; senescence;
XX KW seed germination rate; plant vigor; seedling vigor.
XX OS Arabidopsis thaliana.
XX PN WO2003013227-A2.
XX PD 20-FEB-2003.
XX PF 09-AUG-2002; 2002WO-US025805.
XX PR 09-AUG-2001; 2001US-0310847P.
XX PR 19-NOV-2001; 2001US-0336049P.
XX PR 11-DEC-2001; 2001US-0338692P.
XX PR 14-JUN-2002; 2002US-00171468.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX PA Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
XX PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
XX PI Brown PE;
XX DR WPI: 2003-248221/24.
XX DR P-PSDB; ADD30879.
XX PT New plant transcription factor polynucleotides and polypeptides, useful
XX PT in producing transgenic plants with commercially valuable properties,
XX PT such as an alteration in a plant growth characteristic, e.g. growth rate
XX PT or apomixis.
XX PS Disclosure; SEQ ID NO 907; 454pp; English.
XX CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
XX CC sequences and their encoded proteins which are especially transcription
XX CC factor related cDNA's and proteins. The isolated or recombinant plant
XX CC transcription factor polynucleotides and polypeptides are useful in
XX CC producing transgenic plants with commercially valuable properties, i.e.
XX CC modified or altered desirable traits as compared to a reference plant,
XX CC such as an alteration in a plant growth characteristic, e.g. growth rate,
XX CC germination rate of seeds, vigor of plants and seedlings, or leaf and
XX CC flower senescence. Sequence information related to the polynucleotides
XX CC and polypeptides can also be used in bioinformatic search methods. The
XX CC transgenic plant is useful for growing a progeny plant from a parent
XX CC plant. This sequence represents one of the cDNAs of the invention.
XX SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,85e-83 Length: 1818
Score: 1047.50 Matches: 262
Percent Similarity: 46.94% Conservative: 70
Best Local Similarity: 37.08% Mismatches: 110
Query Match: 33.58% Indels: 265
DB: 9 Gaps: 23

US-09-980-364-2 (1-579) x ADD30878 (1-1818)
Qy 2 AsnAsnAsnTrpLeuGlyPhe-----SerLeuSerProTyrGlu 14
Db 65 TCTAACAACTGGCTTGGCTTTTCCTCTTTCACGGAACAACACTTCTTTCCTCTCATGAA 124
Qy 15 GlnAsnHisHisArgLysAspValTyrSerThrThrThrValValAspValAla 34
Db 125 TACAC----- 130
Qy 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54
```



```

KW seed germination rate; plant vigor; seedling vigor.
XX Arabidopsis thaliana.
XX WO2003013227-A2.
XX 20-FEB-2003.
XX 09-AUG-2002; 2002WO-US025805.
XX 09-AUG-2001; 2001US-0310847P.
XX 15-NOV-2001; 2001US-0336049P.
XX 11-DEC-2001; 2001US-0338692P.
XX 14-JUN-2002; 2002US-00171468.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
XX Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
XX Broun PE;
XX WPI; 2003-248221/24.
XX P-PSDB; ADD30823.
XX New plant transcription factor polynucleotides and polypeptides, useful
XX in producing transgenic plants with commercially valuable properties,
XX such as an alteration in a plant growth characteristic, e.g. growth rate
XX or apomixis.
XX Disclosure; SEQ ID NO 851; 454bp; English.
XX The invention relates to a number of isolated Arabidopsis thaliana cDNA
XX sequences and their encoded proteins which are especially transcription
XX factor related cDNA's and proteins. The isolated or recombinant plant
XX transcription factor polynucleotides and polypeptides are useful in
XX producing transgenic plants with commercially valuable properties, i.e.
XX modified or altered desirable traits as compared to a reference plant,
XX such as an alteration in a plant growth characteristic, e.g. growth rate,
XX germination rate of seeds, vigor of plants and seedlings, or leaf and
XX flower senescence. Sequence information related to the polynucleotides
XX and polypeptides can also be used in bioinformatic search methods. The
XX transgenic plant is useful for growing a progeny plant from a parent
XX plant. This sequence represents one of the cDNAs of the invention.
XX SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.85e-83 Length: 1818
Score: 1047.50 Matches: 262
Percent Similarity: 46.96% Conservative: 70
Best Local Similarity: 37.06% Mismatches: 110
Query Match: 33.58% Indels: 265
DB: 9 Gaps: 23

US-09-980-364-2 (1-579) x ADD30822 (1-1818)
Qy 2 AsnAsnAsnTrpLeuGlyPhe-----SerLeuSerProTyrGlu 14
Db 65 TCTAACAACTGGCTGGCTTTCCTTTCCACGAAACAACCTCTCTTTGCTCTCTCATGAA 124
Qy 15 GlnAsnHisAsArgLysAspValTyrSerSerThrThrThrThrValValAspValAla 34
Db 125 TACAAAC----- 130
Qy 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54
Db 130 ----- 130
Qy 55 PheProSerProPheGlyValValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
Db 131 -----CTGGCTTGGTCAGCGCCATATG-----GACAACTTTTCAACAA 172
Qy 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94

```


Db 278 TCACACACCTAGTAGCTTACACGACTCAGACTACTTCTCCATACCAAT----- 328
Qy 131 GlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgGlnProVal----- 147
Db 329 -----AGCTTGATGCTAGCTCCCAACACGATGCTGTTGTAGCAGCT 373
Qy 148 -----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
Db 374 TGTGACTCCATFACTCTTAACAACAGTAGCTATCATGAGCTTCAAGAGAGTGTCCACAAT 433
Qy 159 AlAlysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
Db 434 CTACAGTCACTTACTTCTTCTCCATGGGACACC-----GCT 469
Qy 179 AsnAsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu-- 194
Db 470 GGTATAATGTTGTAGACAAAGCTTACCATCCGAGACACCGGGGATACGCTAGCGGT 529
Qy 195 -----AlaThrProLysLysThrIleGluSerPheGlyGln 206
Db 530 GGAGCAGCTAGCGGTGTTGAGACGGCCACGCCAAGAGCGTGCATTGGGACACTTTCGGACAA 589
Qy 207 ArgThrSerIleTyArgGlyValThrArgHisArgTrpThrGlyArgTrpGluAlaHis 226
Db 590 CGAACCTCGATCTATCGTGTGTACAAAGACATCGATGGACTGTGCTATATGAGGCTCAT 649
Qy 227 LeuTrpAspAsnSerCysLysArgGlyGlnThrArgLysGlyArgGlnValTyLeu 246
Db 650 CTATGGGATATAGTTGTAGAGGGAAGCGGAGTCTAGGAAGAGGAAGAGTTTACTTG 709
Qy 247 GlyGlyTyAspLysGluGlyAlaAlaArgAlaTyAspLeuAlaAlaLeuLysTy 266
Db 710 GGTGATATGACAAAGAGATAAGCAGCAGACATATGATCTAGCTGCATTAAGTAC 769
Qy 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyArgLysGluValGluGlu 286
Db 770 TGGGTCTCTCAACTACTACTAATTTCCCATTAACAACACGAGAGAGAGTAGAGGAA 829
Qy 287 MetLysHisMetThrArgGlnGluTyValAlaSerLeuArgArgLysSerSerGlyPhe 306
Db 830 ATGAGACACATGACGAGACAGAGTTCGTGCTCCATTAAGAGAAAGTAGTGGATT 899
Qy 307 SerArgGlyAlaSerIleTyArgGlyValThrArgHisGlnHisGlyArgTrpGln 326
Db 890 TCGAGAGCGCTTCGATATCGAGGAGTTACAGGCGATCACCAACATGGAAGATGGCAA 949
Qy 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyLeuGlyThrPheGlyThrGln 346
Db 950 GCAAGGATCGCCGAGTCCCGGAAACAAAGACCTTACTTGGAACTTTTAGCAGTG 1009
Qy 347 GluGluAlaAlaGluAlaTyAspIleAlaAlaLysPheArgGlyLeuThrAlaVal 366
Db 1010 GAAGAAGCAGCAGAGCTTACGATATAGCTGCAATAAGTTTAGAGGACTTAATGCGAGT 1069
Qy 367 ThrAsnPheAspMetAsnArgTyAsnValLysAlaIleLeuGluSerProSerLeuPro 386
Db 1070 ACCACTTCGAGATCAACCGTAGCAGCTGAAGCCATCTAGAGAGTACACTCTTCCC 1129
Qy 387 Ile-----GlySerAlaAlaLysArgLeuLysGluAla----- 397
Db 1130 ATCGAGGAGGCGGAGCTTAAACGGCTCAAGAGCTCAAGCTCTTCAAGGAA 1189
Qy 397 ----- 397
Db 1190 CGCGAGCGGAGATGATAGACCTTGGTTCAAGTTTCCAGTACGGTGGCTCGAGACA 1249
Qy 398 -----AsnArg 399
Db 1250 GGCCTGCTCCACTCATCAAGACTTCAGCTTCAACCTTACCTCTAAGCATTCACAA 1309
Qy 400 ProValProSerMetMetMetIleSerAsnValSerGluSerGluAsnValSer 419
||||| : : : : : |||||

Db 1310 CCATTAGAGCCTTTTCTATCTCTTCAGACAATGACATCTCTCATTACAAACAACAAT 1369
Qy 420 GlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeu---LeuHis 438
Db 1370 GCTCAGATTCCTCTCTTTTAAATCACCATAGTATATCCAGACAACACTTCATCCAC 1429
Qy 439 GlnHis-----GlnGluArgTyAsn 445
Db 1430 CAACAGACCAACAATTTACTTGCAGCAACAGTCGAGCCAGAACTCTCAGCAGCTCTACAAT 1489
Qy 446 GlyTyTrpTyTrp----- 449
Db 1490 GCGTATCTTTCAGCAATCCGGCTCTCTCTCATGAGCTTCTCTCTACCTCTATCGTTGAC 1549
Qy 450 -----AsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheLysGln 465
Db 1550 AACATAATAACAATGAGGCTCTAGTGGGAGCTTACAACACTCAGCA----- 1597
Qy 466 GluAspAspGlnHisPheLeuSerAsn----- 475
Db 1598 -----TTTCTTGGGAACACAGTATTGTTGGTCCAGCTCGACT 1639
Qy 476 -----ThrGlnSerLeuMetThr---AsnIleAspHisGlnSerSerValSerAsp 491
Db 1640 GTTGATTCGACCGAGAGTTTCCAAACCGTTAAACACAGATTACGATATGCTTCCAGTGAT 1699
Qy 492 AspSerValThrValCysGlyAsnValValGlyTyGlyGlyTyGlnGlyPheAlaAla 511
Db 1700 -----GGAACCGGAGGTATAGTGTGG----- 1723
Qy 512 ProValAsnCysAspAlaTyAlaAlaSerGluPheAspTyAsnAlaArgAsnHisTy 531
Db 1723 ----- 1723
Qy 532 TyrPheAlaGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMet 551
Db 1724 -----ACCAGTGAGTCTGTTCCAGGGTCAACACCT----- 1753
Qy 552 ThrAsnAsnValGlySerAsnMetTyTrpHisGlyGluGlyGluValAlaPro 571
Db 1754 -----GGTGT----- 1759
Qy 572 ThrPheThrValTrpAsnAsp 578
Db 1760 GTTTTCTATGTGAATGAG 1780
RESULT 11
ACC00848
ID ACC00848 standard; cDNA; 1935 BP.
XX
XX ACC00848;
AC ACC00848;
DT 16-MAY-2003 (first entry)
XX
DE Gossypium hirsutum oil trait related cDNA sequence SEQ ID NO:486.
XX
KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
transgenic plant; gene; ss.
OS Gossypium hirsutum.
XX
PN WO2003002751-A2.
PD 09-JAN-2003.
XX
PF 27-JUN-2002; 2002WO-US020152.
XX
PR 29-JUN-2001; 2001US-0301913P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.

Db 1350 AATGGTGGGAATCTGAGTCAACGGA-----AGCATTAACATGGTTATTCAT 1400
QY 503 TyrGlyGlyTyrGlnGly----- 508
Db 1401 CAGAGTGGTATGTTGGCGAGGAGCAACAGAGAGAGTGTGCTACTAATTCACCCCA 1460
QY 509 PheAlaAlaProVal-----AsnCysAspAlaTyrAlaAlaSerGluPheAspTyr 525
Db 1461 TTGGCAACGCCAATGTCTTAATAATAGCGATAATATGAGAGCTGTCTTCTGGCTAT 1520
QY 526 AsnAlaAlaArgAsnHisTyrTyrPheAlaGlnGln-----GlnGlnThrGlnGlnSerProGly 544
Db 1521 GGT-----GGTAACATGGGAATCTGCACAGTCGTTTCAATCTACTTATCAAACTGCAAG 1574
QY 545 GlyAspPheProAlaAlaMetThr----- 552
Db 1575 CCAAGTCTCTCAGTGTTCMAACACCAATTTTGGCATGGAATGAATCAGCTTCACATG 1634
QY 553 -----AsnAsnValGlySerAsnMetTyrTyrHisGly-----GluGly 565
Db 1635 GACAATAAATCTGTGGAAAAAATCTG-----TGGGAAGCGCGATGCATTCGAGAGTTGGAAGG 1691
QY 566 Gly-----GlyGluValAlaProThr 572
Db 1692 GGTAGGTTGTTAACTTCTTGTGATAGTGTGAGATGGAATGACA 1736

RESULT 12

ID ACC00847 standard; cDNA; 1959 BP.

XX ACC00847;

XX 16-MAY-2003 (first entry)

XX Catalpa sp oil trait related cDNA sequence SEQ ID NO:484.

XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
XX receptor-like protein kinase; mitogen activated protein kinase; oil;
XX LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
XX CKC-like transcription factor; antisense inhibition; co-suppression;
XX transgenic plant; gene; ss.

XX Catalpa sp.

XX WC003002751-A2.

XX 09-JAN-2003.

XX 27-JUN-2002; 2002WO-US020152.

XX 29-JUN-2001; 2001US-0301913P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX (PION-) PIONEER HI-BRED INT INC.

XX Allen SM, Allen WS, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT,
XX Jones TM, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H, Shen B,
XX Tarczynski MC;

XX WPI; 2003-201509/19.

XX P-PSDB; ABR40853.

XX Novel nucleotide fragment encoding polypeptides having receptor-like
XX protein kinase activity, caleosin-like activity, useful for altering oil
XX phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.

XX Claim 16; Page 501; 542pp; English.

XX The present invention describes an isolated nucleotide fragment (I)
XX comprising a nucleic acid sequence (NS) chosen from a NS encoding a
XX polypeptide (PP) having receptor-like protein kinase activity, mitogen
XX activated protein (MAP)-kinase activity, LIP5-like transcription factor

CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity and CKC-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC0626 to ACC0868 and ABR40891 to
CC ABR40893 represent sequences used in the exemplification of the present
CC invention

XX Sequence 1959 BP; 585 A; 431 C; 450 G; 493 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-89e-79 Length: 1959
Score: 1003.00 Matches: 238
Percent Similarity: 54.20% Conservative: 46
Best Local Similarity: 45.42% Mismatches: 84
Query Match: 32.16% Indels: 156
DB: 7 Gaps: 17

US-09-980-364-2 (1-579) x ACC00847 (1-1959)

QY 2 AsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLysAsp 21
Db 109 CATCAAACTGGCTGGCCTTCTCGCTTTC-----AACCATCT----- 147
QY 22 ValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAspPro 41
Db 148 -----AATCAAGAAAATGGCGGT-----CGTTTCACCCC 159
QY 42 ThrAlaAlaSerAspGluSerSerAlaAlaGlnThrSerPheProSerProPheGlyVal 61
Db 160 TCGTTCAACTTAATCTCTCTCTCCAGCTGC-----GGGGTT 195
QY 62 ValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsnGly 81
Db 196 GTG-----AATCAAGAAAATGGCGGT-----GCGGGT 222
QY 82 CysAlaCysAsnAsnIleHisAsnAsp-----GluGlnAspGlyProLysLeuGluAsn 99
Db 223 GGTGCCGCAACAGATCATCATCTCGGTCTTCCAGTGGAGGCCCAAGCTGAGGAT 282
QY 100 PheLeuGlyArgThrThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
Db 283 TTTCTGGCGGCTCTCTGT-----GGTTCTTCC 309
QY 120 SerGlyCysTyrGlyGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLys 139
Db 310 GCGGACATACGACTCAGTACGACGCGGAGCTCAGGAGTTAATATATCAGATGTAGAG 369
QY 140 ThrTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaAla 159
Db 370 ATATAT-----GGTGACTCGGACTG 390
QY 160 LysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsn 179
Db 391 AAGACAATAGTGTAGTTTCTTACGCGGGTGTTCGATGAACGAGATGATGATCCAG 450
QY 180 AsnAsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrPro----- 197
Db 451 AAGCAGACT-----GCTGTGGCAGCGCTCTGAGCCG 483
QY 198 -----LysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThr 215
Db 484 CCGGCGAAGAAGGCTGCAGAAACATTTGGCCACGCTACCTCGATTTTTCAGAGAGTTACT 543
QY 216 ArgHisArgTptThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGlu 235
Db 544 AGACATAGTGGCTGGAAGATATGAAGCTCATTTTGTGGGACATACTTCCAGAGAGAA 603

Db 1645 -----CCTCAAGCCCAATGAGTTGGAGTCAGTGGAGC 1679
Qy 523 eAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGlnGlnSerPr 543
Db 1680 AGATTAT-----GGACATGGCTACTTCACCCCTGTATGGACCAAAATATGATGATG 1730
Qy 543 oGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyrTyrHisG 563
Db 1731 GACAATGAAATGACACATGACACCAACAGCTTGGAAAT----- 1773
Qy 563 yGluGlyGlyGlyGluValAlaProThrPheThrValTyrPheAsp 578
Db 1774 -TTGGATTGGTTAATCAAGTTCCTCCATGTTTGCTTTGGTGAATGAA 1818
RESULT 15
ACC00799
ID ACC00799 standard; cDNA; 1874 BP.
XX AC ACC00799;
XX 16-MAY-2003 (first entry)
DT Glycine max oil trait related cDNA sequence SEQ ID NO:347.
XX DE
XX KW plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
KW transgenic plant; gene; ss.
XX OS
XX Glycine max.
XX WO2003002751-A2.
XX 09-JAN-2003.
XX 27-JUN-2002; 2002WO-US020152.
XX 29-JUN-2001; 2001US-0301913P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX Allen SM, Allen WB, Epelbaum S, Famodu OO, Harvell LT;
XX Jones TJ, Kinney AU, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
XX Tarczynski MC;
XX WPI: 2003-201509/19.
XX P-PSDB; ABR40764.
XX Novel nucleotide fragment encoding polypeptides having receptor-like
XX protein kinase activity, caleosin-like activity, useful for altering oil
XX phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX Claim 18; Page 364-365; 542pp; English.
XX The present invention describes an isolated nucleotide fragment (I)
XX comprising a nucleic acid sequence (NS) chosen from a NS encoding a
XX polypeptide (PP) having receptor-like protein kinase activity, mitogen
XX activated protein (MAP)-kinase activity, Lip15-like transcription factor
XX activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
XX activity and CKC-like transcription factor activity. Also described: (1)
XX complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
XX (II), operably linked to a regulatory sequence; (3) a plant (IV)
XX comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
XX oil obtained from (V). (I) or its part can be used in antisense
XX inhibition or co-suppression in a transformed plant. (III) is useful for
XX altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
XX canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
XX creating transgenic plants having altered lipid profiles. (I) can also be
XX used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
XX ABR40879 represent sequences used in the exemplification of the present

CC invention
XX Sequence 1874 BP; 542 A; 472 C; 412 G; 448 T; 0 U; 0 Other;
SQ Alignment Scores:
Pred. No.: 4.28e-78 Length: 1874
Score: 987.50 Matches: 234
Percent Similarity: 61.64% Conservative: 60
Best Local Similarity: 49.06% Mismatches: 107
Query Match: 31.66% Indels: 78
DB: 7 Gaps: 13
US-09-980-364-2 (1-579) x ACC00799 (1-1874)
Qy 4 AsnTyrLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLysAspValTyr 23
Db 184 AACTGGCTTTGGTTCTCTCTCTCCCAATGGAATGCTCCGACCTCCGACCTCAGTTC 243
Qy 24 SerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp---ProThr 42
Db 244 CTTCAA-----TACGACGCGCGCTTCC 264
Qy 43 AlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGlyValVal 62
Db 265 GCTACTTCTCATCATCACTACTACTCTGCACACTGTGTACACCAAGGGTGGGC----- 318
Qy 63 ValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTyrPheAsnGlyCys 82
Db 319 -----AACGGAGCCTCAAGTTTGAGCAGATCTCAACACACAGT 357
Qy 83 AlaCysAsnAsnIleHisAsnAspGluGlnAsp-----GlyPro 95
Db 358 GACGTGAGTTTGTGTAATCTTCGTGCGCAGACGTCGCGCCACGTCCCGCGCGCGCGCG 417
Qy 96 LysLeuGluAsnPheLeuGlyArgThr----- 104
Db 418 AAGCTGGAGGATTTCTCGGCGACTCTCCGCGGTGATCGGTACTCCGACAGCCAGACG 477
Qy 105 -----ThrThrIleTyrAsnThrAsnGluAsnValGlyAspGly 117
Db 478 GAGACGCGAGGACTCGCTGACGACATCTACGACACACACACACACACACACACAC 537
Qy 118 SerGlySerGlyCysTyrGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMet 137
Db 538 CACGGTTCTACTTCGTACTTCTCGTGGTGACCA---GCAGGATCTCAAGGCCATTACTGG 593
Qy 138 IleLysThrTyrLeuArgAsnGlnProValAspAsnVal-----AspAsnGln 153
Db 594 ATTCAAGCTTTT-CGACCAACTCCGTTCCGAGGTTGATGATTCGATCCATCCATCGGAA 652
Qy 154 GluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCysasp 173
Db 653 AAGCGCGCCAGCAGGATTCGGGATCTACTCTATTGAGTCTCCGCGACAGATTCGCGCG 712
Qy 174 AsnAsnAsnAspSer-AsnAsnAsnValValAlaGlnGlyLysThr---Ile-AspAspS 192
Db 713 CGTTCTCCGTTGGCCACACCGGAACCTTGTCTCGCTCGCGTTGCTGACTGAGCTCCGAGAAG 772
Qy 192 exValGluAla-----ThrProLysLysThrIleGluSerPheGlyGlnArgT 208
Db 773 CGGTGTCGCGCGGAGTGTCCTAATAGCTCGAAGAAATGCTGGATACCTTCGCCACCGGA 832
Qy 208 hrSerIleTyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuT 228
Db 833 CTTCTATTACAGAGGTGTTACTAGGCACCGATGACAGGAAGATATGAAGCGCATCTAT 892
Qy 228 rAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyG 248
Db 893 GGCACAATAGTTCCAGAGGGGGTTCAGCCAGAAAAGGGGCTCAAGTTTATTGGGGTG 952
Qy 248 lYtYrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTipG 268
Db 953 GATATGATAGGAAGAAAGAAAGCGCGCGAGAGCTTATGATTGGCAGCTCTAAAGTACTGG 1012

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 15:02:27 ; Search time 124 Seconds
(without alignments)
2591.263 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119

Sequence: 1 MNXNLGFLSLPYEQNHHRK.....YHGGGGGVAFTFTVNDN 579

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODELS=frame_p2n.model -DRV=xlh
-Q/cgn2_1/USPTO.spool/US0980364/runat.15072004.095237.25896/app.query.fasta_1.775
-DB=issued_patents NA -QFMT=fastep -SUFFIX=rni -XNMATCH=0.1 -LOEPL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0980364.CGN 1.1 69 @runat.15072004.095237.25896 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOFTERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	913.5	29.3	1738	4	US-09-227-421-4
2	913.5	29.3	1738	4	US-09-479-855-4
3	891	28.6	2148	4	US-09-227-421-1
4	891	28.6	2148	4	US-09-479-855-1
5	461.5	14.8	1680	2	US-08-700-152A-3
6	262	8.4	301	4	US-09-313-294A-6727
7	240.5	7.7	11721	4	US-09-026-039-3
8	240.5	7.7	11721	4	US-09-026-039-3
9	179	5.7	294	4	US-09-313-294A-6087
10	155	5.0	751	4	US-09-533-029-81
11	154.5	5.0	1265	4	US-09-964-850-3
12	152.5	4.9	1577	4	US-09-533-029-11
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 6727, Ap
					Sequence 3, Appli
					Sequence 3, Appli
					Sequence 6087, Ap
					Sequence 81, Appli
					Sequence 3, Appli
					Sequence 11, Appli

13	150	4.8	1055	4	US-09-533-029-21	Sequence 21, Appli
14	149.5	4.8	987	4	US-09-964-850-5	Sequence 5, Appli
15	145	4.6	1405	4	US-09-202-161B-6	Sequence 6, Appli
16	144.5	4.6	921	4	US-09-964-850-1	Sequence 1, Appli
17	141.5	4.5	1132	3	US-08-894-731-3	Sequence 3, Appli
18	140	4.5	1669	4	US-09-026-039-1	Sequence 1, Appli
19	138.5	4.4	806	4	US-09-198-119C-78	Sequence 78, Appli
20	135	4.3	651	4	US-09-198-119C-14	Sequence 14, Appli
21	135	4.3	933	4	US-09-301-666A-1	Sequence 1, Appli
22	135	4.3	933	4	US-09-301-217-1	Sequence 1, Appli
23	134.5	4.3	563	4	US-09-198-119C-64	Sequence 64, Appli
24	132.5	4.2	1281	4	US-09-533-029-65	Sequence 65, Appli
25	132	4.2	913	4	US-09-533-029-33	Sequence 33, Appli
26	131	4.2	577	4	US-09-198-119C-38	Sequence 38, Appli
27	129.5	4.2	1500	3	US-09-300-672-1	Sequence 1, Appli
28	128.5	4.1	174	4	US-09-202-161B-21	Sequence 21, Appli
29	127.5	4.1	761	4	US-09-202-161B-5	Sequence 5, Appli
30	126.5	4.1	651	4	US-09-198-119C-12	Sequence 12, Appli
31	126.5	4.1	944	4	US-09-301-666A-7	Sequence 7, Appli
32	126.5	4.1	944	4	US-09-301-217-7	Sequence 7, Appli
33	125	4.0	349	4	US-09-198-119C-94	Sequence 94, Appli
34	125	4.0	877	4	US-09-198-119C-72	Sequence 72, Appli
35	125	4.0	2161	4	US-09-336-115C-17	Sequence 17, Appli
36	124.5	4.0	884	4	US-09-198-119C-76	Sequence 76, Appli
37	124.5	4.0	2270	4	US-09-336-115C-13	Sequence 13, Appli
38	124	4.0	905	2	US-08-949-603-1	Sequence 1, Appli
39	124	4.0	905	2	US-08-706-270A-1	Sequence 1, Appli
40	124	4.0	905	2	US-08-949-580-1	Sequence 1, Appli
41	124	4.0	905	2	US-08-950-172A-1	Sequence 1, Appli
42	124	4.0	905	4	US-09-198-119C-1	Sequence 1, Appli
43	124	4.0	937	4	US-09-301-666A-5	Sequence 5, Appli
44	124	4.0	937	4	US-09-301-217-5	Sequence 5, Appli
45	124	4.0	964	4	US-09-533-029-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1

US-09-227-421-4
; Sequence 4, Application US/09227421
; Patent No. 6559357
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
; FILE REFERENCE: 023070-090700PC
; CURRENT APPLICATION NUMBER: US/09/227,421
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 09/227,421
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1647)
; OTHER INFORMATION: BANT, ANT ortholog from Brassica napus (Canola)
US-09-227-421-4

Alignment Scores:
Pred. No.: 5.7e-88
Score: 913.50
Percent Similarity: 53.00%
Best Local Similarity: 41.97%
Query Match: 29.25%
DB: 4
Length: 1738
Matches: 217
Conservative: 57
Mismatches: 112
Indels: 131
Gaps: 14

US-09-980-364-2 (1-579) x US-09-227-421-4 (1-1738)

Qy	275	PhaProMetSerGluTyrGluLysGluValGluGluValMetThrArgGlnGlu	294
Db	1018	TTCTCTGTGGAGAAATTATCAGAAGAAGAGATTGATGACATGAAGAACATGACTCGACCAAGAA	1077
Qy	295	TyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArg	314
Db	1078	TATGTTGCTTCACATTGAGAAGAAAACAGTGGTTCCTAGGGGTCTCCATCTATAGA	1137
Qy	315	GlyValThrArgHisGlnHisGlyArgTrrpGlnAlaArgIleGlyArgValAlaGly	334
Db	1138	GGAGTCACCAGACATCACCAGCATGGAAGGTGGCAAGCTCGGATCGGTAGAGTCGCTGGA	1197
Qy	335	AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAsp	354
Db	1198	AACAAGAGATCTCTACCTTGGAACTTCGGAACCTCAAGAGAGACGGCGGAAGCCTATGAT	1257
Qy	355	IleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyr	374
Db	1258	GTAGCAGCTATCAAGTTCCGTGGCACAAACGGCGTGACTTAACCTTTGACATAACAAGGTAC	1317
Qy	375	AsnValLysAlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeu	394
Db	1318	GATGTTGATCGCATAATGGCTAGTAACTCTCTTCTGTGGAGAGATGGCTCAAGG---	1374
Qy	395	LysGluAlaAsnArgProValProSerMetMetIleSerAsnAsnValSerGluSer	414
Db	1375	-----AACAGCAACAGCATCGTGGTCGCAACATTAGCGACGAG	1413
Qy	415	GluAsnSerAlaSerGlyTrrpGlnAsnAlaAlaValGlnHisGlnGly	431
Db	1414	GAAGCCGCTTTAACCCCTCTCGTGAACGGTGGTTCCAATAAGGAAGTGGGT	1464
RESULT 2			
US-09-479-855-4			
; Sequence 4, Application US/09479855			
; Patent No. 6639128			
; GENERAL INFORMATION:			
; APPLICANT: Fischer, Robert L.			
; APPLICANT: Mizukami, Yukiko			
; APPLICANT: The Regents of the University of California			
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility			
; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants			
; FILE REFERENCE: 023070-090720US			
; CURRENT APPLICATION NUMBER: US/09/479,855			
; CURRENT FILING DATE: 2000-01-07			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 4			
; LENGTH: 1738			
; TYPE: DNA			
; ORGANISM: Brassica napus			
; FEATURE:			
; OTHER INFORMATION: canola AINTEGUMENTA (ANT) partial cDNA including			
; OTHER INFORMATION: coding region			
; NAME/KEY: CDS			
; LOCATION: (1)..(1647)			
; OTHER INFORMATION: canola AINTEGUMENTA (ANT)			
US-09-479-855-4			
Alignment Scores:			
Pred. No.: 5,7e-88 Length: 1738			
Score: 913.50 Matches: 217			
Percent Similarity: 53.00% Conservative: 57			
Best Local Similarity: 41.97% Mismatches: 112			
Query Match: 29.29% Indels: 131			
DB: 4 Gaps: 14			
US-09-980-364-2 (1-579) x US-09-479-855-4 (1-1738)			
Qy	4	AsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArg	19
Db	40	AATTTCGTAGGTTCTCGTTGTCTT-----TCAAAATATCTTGAATAATGGTGGTGA	90


```

Db      1078  TATGTTGCTCACATTGAGAGAAAAACCAAGTGGTTTCTCTAGGGGTGCTTCCATCTATAGA 1137
Qy      315  GlyValThrArgHisGlnHisGlyArgTrpGlnAlaAlaArgIleGlyArgValAlaGly 334
Db      1138  GGAGTCACCAGACATCACCAGCATGGAAGGTGGCAAGCTCGGATCGGTAGATCGCTGGGA 1197
Qy      335  AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAsp 354
Db      1198  AACAAAGATCTCTACCTTGGAACTTTTCGGAACCTCAAGAAAGCGGGGAAGCCATGAT 1257
Qy      355  IleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyr 374
Db      1258  GTAGCAGCTATCAAGTTCCTGGGCACAAACGCGGTGACTTAACTTTGACATAACAAGGTAC 1317
Qy      375  AsnValLysAlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeu 394
Db      1318  GATGTTGATCGCATAAATGGCTAGTAACACTCTCTTGCTCGAGAGATGGCTCGAAGG--- 1374
Qy      395  LysGluAlaAsnArgProValProSerMetMetMetMetSerAsnAsnValSerGluSer 414
Db      1375  -----AACAGCAACAGCATCGTGGTCCGCACATTTAGCGACGAG 1413
Qy      415  GluAsnSerAlaSerGlyTrpGlnAsnAlaAlaValGlnHisGlnGly 431
Db      1414  GAAGCGCGCTTTAAACCGCTGCTGTAACCGTGGTTCCATAAGGAAGTGGGT 1464

RESULT 3
US-09-227-421-1
; Sequence 1, Application US/09227421
; Patent No. 6559357
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; FILE OF INVENTION: and Enhancing Asexual Reproduction in Plants
; FILE REFERENCE: 023070-090700PC
; CURRENT APPLICATION NUMBER: US/09/227,421
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 09/227,421
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2148
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (269)..(1936)
; OTHER INFORMATION: AINTEGUMENTA (ANT)
; NAME/KEY: misc feature
; LOCATION: (1109)..(1339)
; OTHER INFORMATION: encodes first AP2 domain
; NAME/KEY: misc feature
; LOCATION: (1340)..(1414)
; OTHER INFORMATION: encodes linker region
; NAME/KEY: misc feature
; LOCATION: (1415)..(1621)
; OTHER INFORMATION: encodes second AP2 domain
US-09-227-421-1

Alignment Scores:
Pred. No.: 2.05e-85 Length: 2148
Score: 891.00 Matches: 227
Percent Similarity: 50.81% Conservative: 56
Best Local Similarity: 40.75% Mismatches: 122
Query Match: 28.57% Indels: 152
DB: 4 Gaps: 17

US-09-980-364-2 (1-579) x US-09-227-421-1 (1-2148)

```



```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..1488
; OTHER INFORMATION: /product= "APETALA 2 (AP2)"
; US-08-700-152A-3

Alignment Scores:
Pred. No.: 1,538-39 Length: 1680
Score: 461.50 Matches: 114
Percent Similarity: 52.11% Conservative: 59
Best Local Similarity: 34.34% Mismatches: 112
Query Match: 14.80% Indels: 47
DB: 2 Gaps: 8

US-09-980-364-2 (1-579) x US-08-700-152A-3 (1-1680)

QY 198 LysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHis 217
Db 544 AAAAAGAGTCGGCGTGGACCAAGATCAAGAAGTCTCAGTATAGAGGTGTACGTTTAC 603
QY 218 ArgTrpThrGlyArgTyrGluAlaHisIleuTrpAspAsnSerCysLysArgGluGlyGln 237
Db 604 CGCGCTACCGGAAGATGGGAATCTCATATTTGGGAC-----TGT----- 642
QY 238 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArg 257
Db 643 -----GGGAACAAGTTTACTTAGTGGATTGACACTGCTCATGCGAGCAGCTCGA 693
QY 258 AlaTyrAspLeuAlaLeuLysTyrTrpGlyThrThrThrThrThrAsnPheProMet 277
Db 694 GCATATGATAGAGCTGCTATTAAATTCGCTGGAGTAGAAGCGGATATCAATTTCAACATC 753
QY 278 SerGluTyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAla 297
Db 754 GAGGATTATGATGATGCTTGAACAGATGACTAATTTAACCAAGGAAGATTCGTACAC 813
QY 298 SerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 317
Db 814 GTACTTCGCGCAAAAGCACAGGCTTCCTCGAGGAAGTTCGAAGTATAGAGGTGTCACT 873
QY 318 ArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAsp 337
Db 874 TTG---CATAAAGTGGTGGTGGGAGTCTGATGGGTCAATTTCTTAGGCAAAAGTAT 930
QY 338 LeuTyrLeuGlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAla 357
Db 931 GTTTATTGGGTTGTTCACACCGAGTCCGAGCTGCTAGAGCTTACGATAAAGCTGCA 990
QY 358 IleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLys 377
Db 991 ATCAATGTAAACGGCAAAAGACCGCTGACCAACTTGTATCCGAGTATTTACGATGAGGAA 1050
QY 378 AlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAla 397
Db 1051 CTCATTCGCGAGTCA-----TCAGGG 1071
QY 398 AsnArgProValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSer 417
Db 1072 AATCTACTACTCCCAACAAGATCACAACTCGATCTGAGCTTGGGAAATTCGGCTAATTCG 1131
QY 418 AlaSerGlyTrpGlnAsnAlaAlaValGlnHisGln----- 430
Db 1132 AAGCATAAAGTCAGATATCGCGCTCAGGATGACCAACCAACCAACAGATTTCTCTCCAC 1191
QY 431 -----GlyValAspLeuSerLeuLeuHisGlnHisGlnGluArgTyrAsn 445
;

```

RESULT 7

US-09-026-039-3
 ; Sequence 3, Application US/09026039
 ; Patent No. 6329567
 ; GENERAL INFORMATION:
 ; APPLICANT: Jofuku, K. Diane
 ; APPLICANT: Okamoto, Jack K.
 ; TITLE OF INVENTION: Methods for Improving Seeds
 ; NUMBER OF SEQUENCES: 103
 ; CORRESPONDENCE ADDRESS:
 ; STREET: Townsend and Townsend and Crew LLP
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM: disk
 ; MEDIUM TYPE: Floppy
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/026,039
 ; FILING DATE: 19-FEB-1998
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/912,272
 ; FILING DATE: 15-AUG-1997
 ; PRIOR APPLICATION DATA: US 08/879,827
 ; FILING DATE: 20-JUN-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/700,152
 ; FILING DATE: 20-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bastian, Kevin L.
 ; REGISTRATION NUMBER: 34,774
 ; REFERENCE/DOCKET NUMBER: 023070-067230US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 11721 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: -
 ; LOCATION: 1..11721
 ; OTHER INFORMATION: /note= "Arabidopsis APETALA2 (AP2)"
 ; OTHER INFORMATION: Genomic sequence"
 US-09-026-039-3

Alignment Scores:
 Pred. No.: 1.56e-14 Length: 11721
 Score: 240.50 Matches: 129
 Percent Similarity: 32.29% Conservative: 57
 Best Local Similarity: 22.40% Mismatches: 115
 Query Match: 7.71% Indels: 276
 DB: 4 Gaps: 19

US-09-980-364-2 (1-579) x US-09-026-039-3 (1-11721)

Qy 7 GlyPheSer-----LeuSerProTyrGluGlnAsnHis----- 17
 Db 995 GGTTCCTCTATTAGCTTCTAACTTGGAGAGATCATACCAGAGGATTGAAGTTTGA 1054
 Qy 18 -----HisArgLysAspValTyrSerSerThr 26
 Db 1055 CCTTCAAGATCAAAATCAAGAAACCAAAAAAATGTGGGATCTAAACGA 1114

Qy 27 ThrThrThrValValAspValAlaGlyGluTyrCysTyrAspPro-ThrAlaAlaSerAs 46
 Db 1115 CGCACCA-----CCAAACACAAAGAGAAGA 1141
 Qy 46 pGluSerSerAlaIleGlnThrSerPheProSerProPheGlyValValValAspAlaPh 66
 Db 1142 AGAATCTGAAGAGTTTCTTATTCTTCACCAAGTAA-----CGGGTTGGATCTTT 1192
 Qy 56 eThrArgAspAsnAsnSer-----HisSerArgAspTrpAspIleAs 80
 Db 1193 CTCTAATCTAGCTCTCTAGCTGTGTATCGAAGATGGATCCGATGACGATGAACCTTAA 1252
 Qy 80 nGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPh 100
 Db 1253 CCGGTCAGACCCCAATAACCCACTGTGCACCATCAGTTCTCTCCCTGAGATGGATTCT-- 1310
 Qy 100 eLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySe 120
 Db 1310 ----- 1310
 Qy 120 rGlyCysTyrGlyGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLysTh 140
 Db 1311 -----AACGGCGGTGGTGTCTTCTGGCTTCTCTCGGGCTCAC-- 1349
 Qy 140 rTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLyl 160
 Db 1350 -TGGTTT----- 1355
 Qy 160 sGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsnAs 180
 Db 1356 -GGTGTAAAGTTTGTCTAGTCGGATCTAGCCACCGGATCGTCGGCGGTAAGCTACCAA 1414
 Qy 180 nAsnValValAlaGlnGlyLysThrIleAspAspSerValGlu---AlaThrPro---Ly 198
 Db 1415 CGTTCGCCGCTGCC-----GTAGTGAGCGCGCACACCGCTTGA 1453
 Qy 198 sLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisAr 218
 Db 1454 AAAGATCGCGGTGGACCAAGATCAAGAGTCTCAGTATAGAGGTGTACGTTTACCG 1513
 Qy 218 qTrpThrGlyArgTyrGluAlaHisLeuTrp-AspAsn----- 230
 Db 1514 CGGTACCGAAGATGGGATCTCATATTTGGTAATATCTCATATTTTAAATTCGTAA 1573
 Qy 231 -----SerCysLysArgGluGln- 237
 Db 1574 TCGATCGTACTTTAGATTATAAATTTAAGTTTTTTTTTTTGTCTCTCTGAATTCAGG 1633
 Qy 238 -----ThrArgLysGlyArgGlnValTyr----- 245
 Db 1634 GACTGTGGAAACAAGTTTACTTAGTAAATTTTATTTTCTCATGTTTTTTTTTGTATT 1693
 Qy 246 -----LeuGlyGlyTyrA 250
 Db 1694 TGGTGTTCAAAAATGTCATCATATAATTTTAAATTATTATTAACTCTGAATAGTGGATTG 1753
 Qy 250 sPlysGluGluLysAlaAla----- 256
 Db 1754 ACATGCTCATGCAGCAGC-TCGGTATTTTTCTCTCTTTGACTCTCTCTATATTGAGTTG 1812
 Qy 256 ----- 256
 Db 1813 TTATTTATTATTTTAAAAATACCGGAAGAAATTTATAAAATTAATTTAATTTT 1872
 Qy 257 -----ArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrThr 273
 Db 1873 TTTTATTTAATAGACATATGATAGAGCTGTATTAATTCCTGGGATAGAACGGATA 1932
 Qy 273 hrAsnPheProMetSerGluTyrGluLysGluValGluGlu----- 286
 Db 1933 TCAATTTCAACATCGAAGATTATGATGATGATCTGAAACAGGTAAATAAATTATAAAC 1992

```
QY 286 ----- 286
Db 1993 TATATTGGTTTATTACAGATTATAAGGTTGGGAGATTATAATTGAAATTGAATTT 2052
QY 287 -----MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerG 305
Db 2053 TATAGATCACTAAATTACCAAGAAGAGTTGCTGACACGTTCTTCGCCGCAAGAAGCACAG 2112
QY 305 LYPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyValArg 325
Db 2113 GCTTCCCTCGAGGAAGTTCGAAGATAGAGGTGTCACCTTG--CAATAAGTGGTGGCTT 2169
QY 325 rpGlnAlaArgIleGlyArgValAlaGly----- 334
Db 2170 GGAAGCTCGAATGGTCAATTCTTAGCCAAAAGTATAATTCTCTCATTTTATATCA 2229
QY 334 ----- 334
Db 2230 CTCGAAACTTCATTTTAGTTTGTATTATTAACCTTGAGTTTGTGTTCTTGAATCTTA 2289
QY 335 ---AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAla----- 350
Db 2290 TAAATAGGTATGTTTATTGGTTTGTTCACACCGAGGTGCAAGCTGCTAGGTAATG 2349
QY 350 ----- 350
Db 2350 TCTTTTGTGTTGATTCTCAACACACACATGTTGTATATGTTTGTGTTTCTCGTACTAATT 2409
QY 351 -----GluAlaTyrAspIleAlaAlaIleLysPheArg 361
Db 2410 GATTTTCATTTTATATATAATACACAGACCTTACGATAAGCTGCAATCAATGTAAC 2469
QY 362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsn 375
Db 2470 GGCAAGAAGCGCGTGACCAACTTGTATCCGAGTATTTACGAT 2511

RESULT 8
US-09-026-039-3/c
; Sequence 3, Application US/09026039
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,039
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,272
; FILING DATE: 15-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/879,827
; FILING DATE: 20-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
```

```
REFERENCE/DOCKET NUMBER: 023070-067230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 11721 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: -
LOCATION: 1..11721
OTHER INFORMATION: /note= "Arabidopsis APETALA2 (AP2)
OTHER INFORMATION: genomic sequence"
US-09-026-039-3

Alignment Scores:
Pred. No.: 1,56e-14 Length: 11721
Score: 240.50 Matches: 129
Percent Similarity: 32.23% Conservative: 57
Best Local Similarity: 22.40% Mismatches: 115
Query Match: 7.71% Indels: 276
DB: 4 Gaps: 19

US-09-980-364-2 (1-579) x US-09-026-039-3 (1-11721)

QY 7 GlyPheSer-----LeuSerProTyrGluGlnAsnHis----- 17
Db 10727 GGTTCCTCTTATTAGTCTTAACCTTCGAGGAGATCATCCAGAGGATTGAAGTTTGA 10668
QY 18 -----HisArgLysAspValTyrSerSerThr 26
Db 10667 CTTCAAAGATCAAATCAAGAAACCAAAAAAATGTGGATCTAAACGA 10608
QY 27 ThrThrValValAspValAlaGlyGluTyrCysTyrAspPro-ThrAlaAlaSerAs 46
Db 10607 CGCACCACA-----CCAAACACAAAGAGAGA 10581
QY 46 pGluSerSerAlaIleGlnThrSerPheProSerProPheGlyValValValAspAlaPh 66
Db 10580 AGAATCTGAAGAGTTTGTATTCTTCACCAAGTAAA-----CGGTTGATCTTT 10530
QY 66 eThrArgAspAsnAsnSer-----HisSerArgAspTrpAspIleAs 80
Db 10529 CTCATAATCTAGCTCTTCAGCTGTTGTTATCGAAGATGATCCGATGACGATTA 10470
QY 80 nGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPh 100
Db 10469 CCGGGTCAGACCAATAACCCACTTGTCCACCATCAGTTCTTCCCTGAGATGGATTCT-- 10412
QY 100 eLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySe 120
Db 10412 ----- 10412
QY 120 rGlyCysTyrGlyGlyAspGlyGlyGlySerLeuGlySerLeuGlyLeuSerMetIleLys 140
Db 10411 -----AACGGCGGTGGTGTGTTCTTCTGGCTTCCTCGGCTCAC-- 10373
QY 140 rTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaAlaLy 160
Db 10372 -TGGTTT----- 10367
QY 160 sGlyLeuSerLeuSerMetAsnSerThrSerCysAspAsnAsnAsnAspSerAsnAs 180
Db 10366 -GGTGTAAAGTTTTCAGTCGATCTAGCCACCGGATCGTCCGGGGTAAAGTACCAA 10308
QY 180 nAsnValValAlaGlnGlyLysThrIleAspSerSerValGlu---AlaThrPro---Ly 198
Db 10307 CGTTCGCCGTGCC-----GTAGTGGAGCCGCGACACCGCTGAA 10269
QY 198 sLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisAr 218
```

Db 10268 AAGAGTCGGCTGGACCAAGATCTCAGATAGAGGTCGTACGTTTACCG 10209
Qy 218 gTTPThrGlyArgTyGluAlaHisLeuTrp-AspAsn----- 230
Db 10208 GCGTACCGGAAGATGGGAATCTCATATTTGGTAATAATCTCPATTTTAAATTTTCGTTAA 10149
Qy 231 -----SerCysLysArgGluGlyGln- 237
Db 10148 TCGATCGTACTTTAGATTAATAAATTTAAAGTTTTTTTGTGTTCTCTCTGATTTTCAGG 10089
Qy 238 -----ThrArgLysGlyArgGlnValTyr----- 245
Db 10088 GACTGTGGAAACAAGTTTACTTTAGCTAAATTTATTTTCTCATGTTTTTTTGTGATTT 10029
Qy 246 -----LeuGlyGlyTyrA 250
Db 10028 TGTGTGTGAAATATGCATCATATAATTTTAAATTTTATTAAATCTCTCAATAGGTGGATTTG 9969
Qy 250 sPLySGluGluLysAlaAla----- 256
Db 9968 ACATGCTCATGCGAGCAGC-TCGGTATTTTCTCTCTCTGACTCTCTCTATATTGATTTG 9910
Qy 256 ----- 256
Db 9909 TTATTTATTTATTTTAAATAATACCGGAAGAAATTTATAAAATTAATTTTAAATTTG 9850
Qy 257 -----ArgAlaTyrAspLeuAlaLeuLysTyrTrpGlyThrThrThrT 273
Db 9849 TTTTATTTAATAGCATATGATAGAGCTGCTATTAAATTCGCTGGAGTAGAGCGGATA 9790
Qy 273 hrAsnPheProMetSerGluTyGluLysGluValGluGlu----- 286
Db 9789 TCAATTTCAACATCGAAGATTATGATGATGACTTGAACAGAGTAAATATATAAATTTATAAAC 9730
Qy 286 ----- 286
Db 9729 TATATGCTGTTTTATTAACGATTTTAAAGGTTTGGAGATTAAATTTGAATTTGAATTT 9670
Qy 287 -----MetLysHisMetArgGlnGluTyrValAlaSerLeuArgArgLysSerSerG 305
Db 9669 TATAGATGACTAAATTTAAACCAAGAGAGTTCGTACACGACTTCGCCGCAAAAGCACAG 9610
Qy 305 LysPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgT 325
Db 9609 GCTTCCTCGAGGAAGTTTGAAGATATAGAGGTGTCACCTTTG---CATAGTGTGTCGTT 9553
Qy 325 rpGlnAlaArgIleGlyArgValAlaGly----- 334
Db 9552 GGGAGCTCGAATGGGTCAATCTTAGCCAAAGATATAATTTCTCTCATTTTATTTCA 9493
Qy 334 ----- 334
Db 9492 CTCGAAAACTTCATTTTAGTTTGTATTTTAACTTTTGTGTTTGTTCGTAATCTTA 9433
Qy 335 ---AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluAlaAla----- 350
Db 9432 TAAATAGGTATGTTTATTTGGTTTTCGACACCGAGTCGAGCTGTAGTAAATG 9373
Qy 350 ----- 350
Db 9372 TCITTTTGTGTTGATTCTACAACACATGCTGTGTATAATGTTTCTCGTTACTAAT 9313
Qy 351 -----GluAlaTyrAspIleAlaAlaIleLysPheArg 361
Db 9312 GATTTTCATTTTATATATAATACACAGAGCTTACGATAAGCTGCAATCAAAATGTAAC 9253
Qy 362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsn 375
Db 9252 GGCAAGACGCGGTGACCAACTTTGATCCGAGTATTACGAT 9211
RESULT 9
US-09-313-294A-6087

; Sequence 6087, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6087
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351215H1
US-09-313-294A-6087

Alignment Scores:
Pred. No.: 1,94e-10 Length: 294
Score: 179.00 Matches: 42
Percent Similarity: 56.47% Conservative: 6
Best Local Similarity: 49.41% Mismatches: 4
Query Match: 5.74% Indels: 33
DB: 4 Gaps: 1

US-09-980-364-2 (1-579) x US-09-313-294A-6087 (1-294)
Qy 251 LysGluGluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyrTrpGlyThrThr 270
Db 5 AAAGAGAGAAAGCTTACCAGGGCTTACAAATTTAGCTGCTCTCAAGTACTGGGGAACCA 64
Qy 271 ThrThrAsnPhePro----- 276
Db 65 AGCATACAGTTTCCAGTATTTGCATATATTTTCAGTCTCAGTAATATATCTT 124
Qy 277 -----MetSe 278
Db 125 TTTTTCCTTTTACCTGCTCTTTGGATATTATACGATTTATTTTTCATGCTAGATCAG 184
Qy 278 rGluTyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSe 298
Db 185 CAACTATGAGAGGACCTTGGAGAGATGATATATATGTGACGACAGAAATATGTTGCTC 244
Qy 298 rLeuArgArgLys 302
Db 245 ACTTAGAAGGAAG 257

RESULT 10
US-09-533-029-81
; Sequence 81, Application US/095333029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22

EARLIER APPLICATION NUMBER: 60/125,814

NUMBER OF SEQ ID NOS: 121

SOFTWARE: PatentIn version 3.0

SEQ ID NO 81

LENGTH: 751

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: G25

OTHER INFORMATION: G25

US-09-533-029-81

Alignment Scores:
Pred. No.: 3,136-07 Length: 751
Score: 155.00 Matches: 50
Percent Similarity: 45.08% Conservative: 37
Best Local Similarity: 25.91% Mismatches: 94
Query Match: 4.97% Indels: 12
DB: 4 Gaps: 4

US-09-980-364-2 (1-579) x US-09-533-029-81 (1-751)

Qy 265 LysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluVal 284
Db 77 AAAATGTGGGGAGCTATCATTTCTGTTTCATCTGGTGAATCTGAGTCAGAACCG 136
Qy 285 GluGluMetLysHisMetThr-----ArgGlnGluTyrValAlaSerLeuArgArgLys 302
Db 137 AGTCAACTCGGCTCTGTGTAGCAGCAGGAAGGTAACCGCTCTCAGTCAGTGAAGAA 196
Qy 303 SerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHis 322
Db 197 AGAGATGGGAACGAGAGAGGAAGAAATCTGTACAGAGGATA---AGCAGAGCCCATGG 253
Qy 323 GlyArgTrpGlnAlaArgGileGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThr 342
Db 254 GGCAATGGCAGCGGAGATCTGTACCCGAGCAAGAGTGTTACGTGTCTGGCTTGGCACA 313
Qy 343 PheGlyThrGlnGluGluAlaGluAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGly 362
Db 314 TTCAAAACCGCGCAGAGCTGCTCGAGCTTACGAGCTGTGCGCATCAAAATCCGTGGC 373
Qy 363 LeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSer 382
Db 374 CGAAAGCCAACTGAATTTCCCA-----AACACTCAAGTAGAGAGAGAGCC 421
Qy 383 ProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgProValPro 402
Db 422 GATATAAACCGGGGGGAATCAAAATGAGCTGATTTCCGAA-----AAC 466
Qy 403 SerMetMetIleSerAsnValSerGluSerGluAsnSerAlaSerGlyTyrGln 422
Db 467 CAAGTAGAGACTATCGAGAGACTGATGTCATTTGAGGATACATGAGATCTATCAG 526
Qy 423 AsnAlaValGlnHisGlnGlyValAspLeuSerLeuHisGlnHisGlnGlu 442
Db 527 ATTCGGTGTGCGCAGCAATCGCGCAGCATATTCGAAATTTATGGAGCTATCAGAC 586
Qy 443 ArgTyrAsnGlyTyrTyrAsnGlyAsnLeuSer 455
Db 587 TCCAAATTAATCTCTTATTTCCCGCGCGTTTGTCTACT 625

RESULT 11

US-09-964-850-3

Sequence 3, Application US/09964850

Patent No. 6441276

GENERAL INFORMATION:

APPLICANT: Ikeda, Yoshihisa

APPLICANT: Chua, Nam-Hai

TITLE OF INVENTION: ESR2 - A PLANT GENE THAT CAN PROMOTE PLANT REGENERATION AND

TITLE OF INVENTION: TRANSFORMATION

FILE REFERENCE: 2312-111

CURRENT APPLICATION NUMBER: US/09/964,850

CURRENT FILING DATE: 2001-09-28

NUMBER OF SEQ ID NOS: 8

SOFTWARE: PatentIn version 3.0

SEQ ID NO 3

LENGTH: 1265

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: CDS

LOCATION: (76)..(1059)

US-09-964-850-3

Alignment Scores:
Pred. No.: 86-07 Length: 1265
Score: 154.50 Matches: 80
Percent Similarity: 38.22% Conservative: 40
Best Local Similarity: 25.48% Mismatches: 111
Query Match: 4.95% Indels: 84
DB: 4 Gaps: 12

US-09-980-364-2 (1-579) x US-09-964-850-3 (1-1265)

Qy 251 LysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyThrThr 270
Db 73 AAAATGGAAGAGCTTGTGAAATCTC-----ACC 102
Qy 271 ThrThrThrAsnPhe-----ProMetSerGluTyrGluLysGlu----- 283
Db 103 GAATACCACTCCACAGACCTAATCTCTCACAATAATCTTCACCTGAACCTACAGCC 162
Qy 284 -----ValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 301
Db 163 TCACCTGTTAGCGGCAACCGCAACTGCTTCAAAAGATACCACTGTACCATCCCGGA 222
Qy 302 LysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGln 321
Db 223 GCTGGCAGCAGCAGCAGCAGG-----TACCGCGGCGTA---CGCGGAGGCGC 267
Qy 322 HisGlyArgTrpGlnAlaArgGileGlyArgValAlaGlyAsnLysAspLeuTyrLeuGly 341
Db 268 TGGGACAGATACCGCGGAGATACGTGACCCCAATGTCAAGAGAGAGACGTGGCTCGGA 327
Qy 342 ThrPheGlyThrGlnGluAlaGluAlaGluAlaTyrAspIleAlaAlaIleLysPheArg 361
Db 328 ACATTTGACAGCGGGAACACCGCTTGTGCTTACGACTCTGCGGCTCGTCTTCTCGT 387
Qy 362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGlu 381
Db 388 GGAGCAAGGCTCGTACTAATTT-----ACTTATCCGACAGCTGTCTATATGCT 438
Qy 382 SerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgProVal 401
Db 439 GAACCAAGTTTTCTTTT---TCCAAAGAAATCTTCGCCGCTCGCTCGTCTCTCT 495
Qy 402 ProSerMetMetIleSerAsn----- 409
Db 496 CCTTCTCTACCGTTAGATTCTCTTACCCAAACTTTTTCGGTGCACCGCAGCAGAGG 555
Qy 410 -----AsnValSerGluSerGluAsnSerAlaSer 419
Db 556 ATCTATAATACAGACTCTATCTTCTTACGCGCGCTGTGTCTCTCTCTCTCTCTCT 615
Qy 420 GlyTrpGlnAsnAla-AlaValGlnHisGlnGlyValAspLeuSerLeuHisGln 439
Db 616 CCTATAATAACTCTTCAACGCGCTCATCAT-----CTTCTTACTCA 657
Qy 439 HisGlnGluArgTyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSer----- 456
Db 658 GCATCGAAACCGGATGGGTTT-TCTTATTCGAAACGAAACACAGAGTCTGTTTTCCC 716
Qy 457 -----GluSerAlaArgAlaCysPheLysGlnGluAspAspGln----- 469
Db 717 GGAAGAATCTTCTGATCTGGTCTATTACAGAGGTCGTTCAAGAGTCTTCTTGAAGAAAA 776


```
Qy 470 -----HisHisPheLeuSe 474
Db 777 TCGGGGGTTCCTCTCTCCACACACACCGCGGTGACTAGCCATCATGACCACTC 836
Qy 474 rAsnThrGlnSerLeuMetThrAsnIleAspHisGlnSerSerValSerAspSerVa 494
Db 837 TGGTATTCTCTGCTCTACT--ATATACTTGAAATATATGTTCAAGAGACTAAGGA 893
Qy 494 lThrValCysGlyAsnValValGlyTyrGlyTyrGln 507
Db 894 GACTTTGCTCGAAACTAGATCGCTACGGGAATTTTCAA 933

RESULT 12
US-09-533-029-11
; Sequence 11, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omalra
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; EARLIER FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1577
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G4
US-09-533-029-11

Alignment Scores:
Pred. No.: 1.85e-06 Length: 1577
Score: 152.50 Matches: 70
Percent Similarity: 42.75% Conservative: 48
Best Local Similarity: 25.36% Mismatches: 110
Query Match: 4.89% Indels: 49
DB: Gaps: 11

US-09-980-364-2 (1-579) x US-09-533-029-11 (1-1577)

Qy 167 AsnSerThrSerCysAspAsnAsnAsnAspSerAsnAsnVal-----182
Db 273 GACTCGGCTTTGACTCGGAAGACGATGATGCTTCTGCAATGTTAAGCCTTCGTC 332
Qy 183 ---ValAlaGlnGlyThrIleAspAspSerValGluAlaThr-----196
Db 333 TTACCCCAACTACTAAGCCGAGTCTCGCTTTCGCTTCCACTGTAGTTCAGCATAT 392
Qy 197 ProLysLysThrIleGluSerPheGlyGln-----ArgThr 208
Db 393 GCCAAGAAACTGTAGAGTCGCTGAGCAAGCTGAGAAATCTTCTAAGAGAGAGGAAG 452
Qy 209 SerIleTyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTrp 228
Db 453 AATCAATACCGAGGATTAGCGACGCTCTTGG---GGAAATGGGTGCGGAGATCCGT 509
```

```
Qy 229 AspAsnSerCysLysArgGluGlnThrArgLysGlyA-GlnValTyrLeuGlyGly 248
Db 510 GAT-----CCGAGAAAGGCTCCCGAGAAATGGCTTGGAAACA 545
Qy 249 TyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGly 268
Db 546 TTCGACATCTGTAGAGACGACAGAGCTTATGATGCTGACGACGCGAATCCGTGTC 605
Qy 269 ThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLys 288
Db 606 ACAGAAAGCTAAGTGAATTTCCCGAGGAGAAACCCCTAGCTGCTATCCACAGAAACGT 665
Qy 289 HisMetThrArgGlnGluTyrValAlaSerLeuArgLysSerSerGlyPheSerArg 308
Db 666 CCTAGTGTAGACAAAT-----AATCTTCAGAAATCAGTGGCTAAACCAACAAA 716
Qy 309 GlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArg 328
Db 717 AGCGTAACCTTGGTTCAGCAGCAACACATCTGAGTCAG-----755
Qy 329 IleGlyArgValAlaGlyAsnLys---AspLeuTyrLeuGlyThrPheGlyThrGlnGlu 347
Db 756 -----CAGTACTGCAACAACCTCCTTTGCACAACTCTTTTGGTGATATGAGTTTCATGAA 809
Qy 348 GluAlaAlaGluAlaTyrAsp-IleAla---AlaIleLysPheArgGlyLeuThrAlaVa 366
Db 810 GAGAGCCTCAGATGTACAAATCAGTTGGTTTAAACAACCTGTTGATGCTGGAGT 869
Qy 366 lThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPr 386
Db 870 AACATGGATACCACTATTTTCAGTTCCGATCAGGCGCACTAATCCTTCGACTGTTCTGAG 929
Qy 386 oileGlySerAlaAlaLysArgLeuLysGluAlaAsnArgProValProSerMetMetWe 406
Db 930 TTCGGTGGAGTGATCAGCGCCCTAAAA-----CACCGAG-ATCTCTTCAATGCT 979
Qy 406 rIleSerAsnAsnValSer-----GluSerGluAsnSerAla 418
Db 980 TGTCAATAACCAACGAAGCATCATTTGTTGAAGAAACCAATGACGCC 1025

RESULT 13
US-09-533-029-21
; Sequence 21, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omalra
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER FILING DATE: 60/125,814
; EARLIER APPLICATION NUMBER: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G19
US-09-533-029-21
```

Alignment Scores: 1.89e-06 Length: 1055
 Pred. No.: 150.00 Matches: 68
 Score: 40.61% Conservative: 38
 Percent Similarity: 26.05% Mismatches: 105
 Best Local Similarity: 4.81% Indels: 50
 Query Match: 4 Gaps: 10
 DB: 4

US-09-980-364-2 (1-579) x US-09-533-029-21 (1-1055)

QY 240 LysGlyArgGlnValTyrLeuGlyTyrAspLysGluGlyLysAlaAlaArgAlaTyr 259
 DB 121 AAGGCCGTAACTACCGCTGAGAACTCTGTCAGAGCTCGATGCTTCGCCGCCGAC 180
 QY 260 AspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPhe---PrometSer 278
 DB 161 GAC-----TTCTGGGTTCTATTCCACCTCCAAACTCCATCCCAACCAAC 225
 QY 279 GluTyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSer 298
 DB 226 CAAGTTAACTGAAAGAGGAGCGCATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 285
 QY 299 LeuArgAlaGlySerSerGlyPheSerArgLysAlaSerileTyrArgGlyValThrArg 318
 DB 286 AGGAAGAGGAG-----AATGTTTATAGGGGATA---CGT 318
 QY 319 HisHisGlnHisGlyArgTyrGlnAlaArgileGlyArgValAlaGlyAsnLysAspLeu 338
 DB 319 AAGCGTCCATGGGAAATGGCGCTGAGATTCGAGATCCAGCAAGAGGTGTAGAGTT 378
 QY 339 TyrLeuGlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaile 358
 DB 379 TGGCTGTGTAGCTTCAACACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 438
 QY 359 LysPheArgGlyLeuThrAlaValThrAsnPhe---AspMetAsnA-gTyrAsnValLys 377
 DB 439 CAGATCCGTGTGTGATAAGCAAGCTCAACTTCCAGATCTGCACCATCTCTCTCTCT 498
 QY 378 AlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAla 397
 DB 499 AATTATATCTCTCCCGCTCATCGCA-----CGATCAACC 534
 QY 398 AsnArgProValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSer 417
 DB 535 GATCAGCT---CGCGCAAGAGGTCTCGCTGCTCTCTCAGAGTGAGAGCGAGTTAAGT 591
 QY 418 AlaSerGlyTyrGlnAsnAlaAlaValGlnHisGlnGlnGlyValAlaAspLeuSerLeu 437
 DB 592 CAGCCGAGTTTCCCGCTGAGGTATAGGATTTGAAATGGGACGAGTTTCAG----- 645
 QY 438 HisGlnHisGlnGluArgTyrAsnGlyTyrTyrTyrTyrAsnGlyGlyAsnLeuSerSerGlu 457
 DB 646 -----AACCTGAGTTACGGA 660
 QY 458 SerAlaAlaGlyPheGlyGlnGluAspAspGlnHisHisPheLeuSerAsnThrGln 477
 DB 661 -----TTTGAGCCGATATATGATCTGAAACAGCAGATATCAGCTTCGAA 705
 QY 478 SerLeuMetThr-----AsnIleAspHisGlnSerSerValSerAspSerVal 494
 DB 706 TCGTCTCTTGAGCTGGAGCGGTAAACAGCGGAGGAGCAACCGAGTTCAGTTGATCGCTT 765
 QY 495 Thr 495
 DB 766 TCC 768

RESULT 14

US-09-964-850-5
 ; Sequence 5, Application US/09964850
 ; Patent No. 6441276
 ; GENERAL INFORMATION:
 ; APPLICANT: Ikeda, Yoshihisa

APPLICANT: Chua, Nam-Hai
 ; TITLE OF INVENTION: EST2 - A PLANT GENE THAT CAN PROMOTE PLANT REGENERATION AND
 ; FILE REFERENCE: TRANSFORMATION
 ; CURRENT APPLICATION NUMBER: US/09/964,850
 ; CURRENT FILING DATE: 2001-09-28
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 987
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(984)
 ; US-09-964-850-5

Alignment Scores: 1.87e-06 Length: 987
 Pred. No.: 149.50 Matches: 79
 Score: 37.17% Conservative: 47
 Percent Similarity: 23.30% Mismatches: 119
 Best Local Similarity: 4.79% Indels: 95
 Query Match: 4 Gaps: 11
 DB: 4

US-09-980-364-2 (1-579) x US-09-964-850-5 (1-987)

QY 253 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThrThr 272
 DB 4 GAAAGAGCTTCAGAACTTC-----ACGAATCT 33
 QY 273 ThrAsnPhe-----PrometSerGluTyrGluLysGlu----- 283
 DB 34 ACCACTCACCAGACCTAATCTCTCACAATAATCTTCACTGAACCTACAGCCTCACT 93
 QY 284 ValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSer 303
 DB 94 GTAGCCGCAACCGCAAACTGTCTTCAAAAGATACCATGTAAACCATCCGCGGAGCTGC 153
 QY 304 SerGlyPheSerArgGlyAlaSerileTyrArgGlyValThrArgHisGlnHisGly 323
 DB 154 AGCAGCACACGAGG-----TACCGCGCGTA---CGCCGAGGCGCGGGGA 198
 QY 324 ArgTyrGlnAlaArgileGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPhe 343
 DB 199 CGATACCGCGCGAGATACGTGACCAATGTGAAAGAGAGAGACGTGTGCTCGGAACATTT 258
 QY 344 GlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeu 363
 DB 259 GACACGGCGCAACAGCCGCTGTGCTTACGACTCTGGGCTCGTCCCTTCGTGGAGCA 318
 QY 364 ThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSer--- 382
 DB 319 AAGGCTCGTACTAATTT-----ACTTATCCGACAGCTGTCTATTATGCTGAACCA 369
 QY 383 -----ProSerLeu 385
 DB 370 AGTTTTCTTTTCCAAAGAAATCTTCGCCGTCTGCTGCTTCTCTCTCTCTCTCTCT 429
 QY 386 ProIleGlySer-----AlaAlaLysArg----- 393
 DB 430 CCGTTAGATTCTCTTACCACAAATTTTACGGTGCACCGGACGAGCAGAGGATCTATAAT 489
 QY 394 -----LeuLysGluAlaAsnArgProValProSerMetMetMetIleSer 408
 DB 490 ACACAGTCTATCTTTCACGACGCGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 549
 QY 409 AsnAsnValSerGluSerGluAsnSerAlaSerGlyTyrGlnAsnAlaAlaValGlnHis 428
 DB 550 AACTCTTTCACGGCTCATCTCTTACTCAGCATCGAAACGCGCATGGTTCCTTCTTAT 609
 QY 429 HisGlnGlyValAsp-LeuSer----- 435

Db 610 TCGAAACGAAACAAACACGAGTCGTTTTTCCCGGAAGAAATCTTCTGATACCTGGTCTATTA 669
QY 436 -----LeuLeuHisGlnHi 440
Db 670 CAAGAGTCGTTCAAGAGTCTTGAAGAAAAATCGCGGCTTCCTCCCTCTCCACCAACA 729
QY 440 sGlnGluArgTyrAsnGlyTyrTyrAsnGlyGly-----AsnLeuSerSerG1 457
Db 730 CCACCCCGCGTG-ACTAGGCATCATACAACTCTGTTATTTCTCTAATCTCCTCATATATA 788
QY 457 uSerAlaArgAlaCysPheLeuGlnGluAspAspGlnHisHisPheLeuSerAsnThrG1 477
Db 789 CICTGAARATATGTTCAAGAGACTAAGGAGACTTTGTCTCGAAACTAGATCGCTACGG 848
QY 477 nSerLeuMetThrAsnIleAspHisGlnSerSerValSerAspAspSerValThrValCy 497
Db 849 GAATTTTCAAGCTAATACGACCGCGCTAAGAGCCGTCGACGCGGTGTATCTCGTTGGG 908
QY 497 sGlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCys 515
Db 909 ATCAACGAGTGGGGGTATCAAGAAATGTTGATGACGGAATCAGTTAGGCTGT 963

RESULT 15

US-09-202-161B-6
; Sequence 6, Application US/09202161B
; Patent No. 6653533
; GENERAL INFORMATION:
; APPLICANT: Purdue Research Foundation
; TITLE OF INVENTION: GENES ENHANCING DISEASE RESISTANCE IN PLANTS
; FILE REFERENCE: 7024-371
; CURRENT APPLICATION NUMBER: US/09/202,161B
; CURRENT FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: PCT/US97/10382
; PRIOR FILING DATE: 1997-06-12
; PRIOR APPLICATION NUMBER: 60/046,494
; PRIOR FILING DATE: 1997-05-14
; PRIOR APPLICATION NUMBER: 60/019,633
; PRIOR FILING DATE: 1996-06-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: ASCII
; SEQ ID NO 6
; LENGTH: 1405
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-09-202-161B-6

Alignment Scores:

Pred. No.:	9.83e-06	Length:	1405
Score:	145.00	Matches:	50
Percent Similarity:	46.55%	Conservative:	31
Best Local Similarity:	28.74%	Mismatches:	67
Query Match:	4.65%	Indels:	26
DB:	4	Gaps:	7

US-09-980-364-2 (1-579) x US-09-202-161B-6 (1-1405)

QY 144 AsnGlnProValAspAsnValAspAsnGlnGlu---AsnGlyAsnAlaAlaLysGlyLeu 162
Db 366 AACAAAGCATGTTTCTGAGCATACGAAATAATCCAAATTCAGATTACAAAGAGTTGTG 425
QY 163 SerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSerAsnAsnVal 182
Db 426 AGGATTATATCTACAGATCCGATGCTACAGATCTTCCTCGATGATGAAGCCGGAATACT 485
QY 183 ValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrProLysLysThrIleGlu 202
Db 486 GTACGGAGATGAAGAGCCACGTCACGGAGATCAACCTTATGCG-----TCAACCAAA 539
QY 203 SerPheGly-----GlnArgThrSerIle----- 210
Db 540 TCGATCGCGCACAGAAACAGAGATCGGTCTCCGGAATTCACGCTCAGCTCGCGAAA 599
QY 211 ---TyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTyrAsp 229

Db 600 AAGTTTATAGAGCGGTTCTGTCAAAGACCGTGG---GGTCGTTGGGCTGCAGAGATTCCGGAC 656
QY 230 AsnSerCysLysArgGlnGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyr 249
Db 657 CCGACC-----CGGGGAAACCGGTGTGTTGGGTACTTAT 692
QY 250 AspLysGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThr 269
Db 693 GACACCCAGAAAGACGCTGCTGTTTACGATAAAGCTCAGTTAAGCTCAAAGGTCCT 752
QY 270 ThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLysHis 289
Db 753 GACGCCGTTACCAATTTTCCGGTATCAACA---ACGCCGAGGTTAACGGTCACGGTTTACG 809
QY 290 MetThrArgGlnGluTyrValAlaSerLeuArgArgLysSer 303
Db 810 GAAACCGAAACCGAGTCTGTTGCCGCGGTGGAGATAAAAGC 851

Search completed: July 19, 2004, 18:30:25
Job time : 164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 17:20:03 ; Search time 742 Seconds
(without alignments)
3806.904 Million cell updates/sec

Title: US-09-980-364-2
Perfect score: 3119
Sequence: 1 MNNNLGFLSPYEQNHRRK.....YYHGEGGCVAPTFVWNND 579

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3190992 seqs, 2439311697 residues

Total number of hits satisfying chosen parameters: 6381984

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09980364/runat_15072004_095239_25984/app.query.fasta_1.775
-DB=published Applications NA -QFMT=fastap -SUPER=rnpb -MINMATCH=0.1
-LOFPCLO=0 -LOFPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdd -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEA=SIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09980364 @CGN 1 1 516 @runat_15072004_095239_25984
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
------------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1
US-10-183-687-351
; Sequence 351, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarzynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01456 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27

1	1310	42.0	2588	13	US-10-183-687-351	Sequence 351, App
2	1296	41.6	2463	13	US-10-183-687-355	Sequence 355, App
3	1119.5	35.9	1990	13	US-10-183-687-361	Sequence 361, App
4	1111.5	35.6	2169	13	US-10-424-599-91337	Sequence 91337, A
5	1110.5	35.6	2088	16	US-10-374-780A-1738	Sequence 1738, App
6	1075.5	34.5	1824	13	US-10-183-687-353	Sequence 353, App
7	1047.5	33.6	1818	13	US-10-225-066A-365	Sequence 365, App
8	1047.5	33.6	1818	13	US-10-225-066A-855	Sequence 855, App
9	1047.5	33.6	1818	13	US-10-225-066A-911	Sequence 911, App
10	1047.5	33.6	1818	16	US-10-374-780A-333	Sequence 333, App
11	1017	32.6	1935	13	US-10-183-687-486	Sequence 486, App
12	1003	32.2	1959	13	US-10-183-687-484	Sequence 484, App
13	1000	32.1	1878	13	US-10-183-687-482	Sequence 482, App
14	996.5	31.9	1954	13	US-10-183-687-490	Sequence 490, App
15	996	31.9	2116	13	US-10-424-599-100930	Sequence 100930, App
16	987.5	31.7	1874	13	US-10-183-687-347	Sequence 347, App
17	986.5	31.6	2128	16	US-10-374-780A-1763	Sequence 1763, App
18	979	31.4	2374	13	US-10-183-687-488	Sequence 488, App
19	977	31.3	2052	13	US-10-225-066A-1013	Sequence 1013, App
20	977	31.3	2052	16	US-10-374-780A-2495	Sequence 2495, App
21	974.5	31.2	1941	13	US-10-225-066A-381	Sequence 381, App
22	974.5	31.2	1941	16	US-10-374-780A-345	Sequence 345, Appli
23	973.5	31.2	1926	14	US-10-024-632-5	Sequence 5, Appli
24	971.5	31.1	1920	17	US-10-437-963-49245	Sequence 49245, A
25	971	31.1	2168	13	US-10-183-687-337	Sequence 337, App
26	964.5	30.9	1758	13	US-10-424-599-70525	Sequence 70525, A
27	960	30.8	2147	13	US-10-183-687-359	Sequence 359, App
28	957.5	30.7	1290	13	US-10-435-114-783	Sequence 783, App
29	956.5	30.7	1985	13	US-10-183-687-333	Sequence 333, App
30	949.5	30.4	1753	13	US-10-425-114-20762	Sequence 20762, A
31	948	30.4	794	13	US-10-425-114-20762	Sequence 14259, A
32	948	30.4	1479	13	US-10-425-114-14859	Sequence 3380, App
33	948	30.4	1864	13	US-10-183-687-331	Sequence 37126, A
34	947.5	30.4	1466	13	US-10-425-114-3280	Sequence 8, Appli
35	944.5	30.3	2102	17	US-10-437-963-37126	Sequence 1, Appli
36	942	30.2	2010	14	US-10-024-632-8	Sequence 20346, A
37	942	30.2	2344	14	US-10-437-963-20346	Sequence 591, App
38	941.5	30.2	1830	17	US-10-225-066A-591	Sequence 399, App
39	941	30.2	1500	16	US-10-374-780A-399	Sequence 19324, A
40	941	30.2	1500	13	US-10-425-114-19324	Sequence 3, Appli
41	935.5	30.0	1182	13	US-10-024-632-3	Sequence 4, Appli
42	924	29.6	2323	15	US-10-059-911-4	Sequence 84937, A
43	913.5	29.3	1738	15	US-10-437-963-84937	
44	901.5	28.9	1869	17		


```

Db      1842 GTTGTTCAGTGATGGTGATCAAAATCCAGAAGCAATCATGGTTTGGTGATAATGAG 1901
QY      526 -----
Db      1902 ATAAAGCACTGGTTATGAAGTGTGTATGCTCTGCAACTGATCTTATCATGCACAT 1961
QY      527 AlaArgAsnHisTyrTyrPheAlaGlnGlnGlnGlnSerProGlyGlyAsp 546
Db      1962 GCAAGGAACCTTGATTATCTTACTCAACAGCAA----- 1994
QY      547 PheProAlaAlaMetThrAsnAsnValGlySerAspMetTyrTyrHisGlyGluGlyGly 566
Db      1995 -----TCATCTTCTCTTATACAGTGAAGGCTAGTGCAATGATCAAGGGTCTGCATGC 2048
QY      567 GlyGluValAlaProThr 572
Db      2049 AATACTTGGGTTCCTCAACT 2066

RESULT 2
US-10-183-687-355
; Sequence 355, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 355
; LENGTH: 2463
; TYPE: DNA
; ORGANISM: Glycine max
US-10-183-687-355

Alignment Scores:
Pred. No.: 7,74e-141 Length: 2463
Score: 1296.00 Matches: 324
Percent Similarity: 54.78% Conservative: 66
Best Local Similarity: 45.51% Mismatches: 146
Query Match: 41.55% Indels: 176
DB: 13 Gaps: 29

US-09-980-364-2 (1-579) x US-10-183-687-355 (1-2463)
QY      4 AsnTrpLeuGlyPheSerLeuSerProTyrGlu-----GlnAsnHisHisArg 19
Db      197 AATTGTGTAGGTTTTCTCTCTCTCTCTCAAGAACACCCCTTCTAGTCAAGATCACTCTCAA 256
QY      20 -----LysAspValTyrSerSerThrThrThrVal 30
Db      257 ACGGCACCTCTCGTTTTGGTTTCAACCTGATGGAATCTCAACACT----- 304
QY      31 ValAspValAlaGlyGlyTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAla 50
Db      305 ---GATGTAGCAGGAGAC---TGCTTTGATCTCACTTCT-----GACTCAACTCCT 349

```

```

QY      51 IleGlnThrSerPheProSerProPheGlyValValValAspAlaPheThrArgAspAsn 70
Db      350 CATTTACTCAACCTTCCTCTCT---TACGGC---ATATACGAAGCTTTTCATAGGAGCAAC 403
QY      71 AsnSerHis---SerArgAspTrpAspIleAsn-----Gly 81
Db      404 AATATTACACCACTCAAGATTGGAAGGAGAACTACACAGCCAAAACCTTGTATTGGGA 463
QY      82 CysAlaCysAsnAsn-----IleHisAsn-----AspGluGlnAspGlyPro 95
Db      464 ACTTCATGCAGCAACCAAAACATGAACCAACCATCAGCAACCAACCAACCAACCAACCA 523
QY      96 LysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsn---Val 114
Db      524 AAGCTTGAATAACTTCTTCGGTGGACACTCA-----TTTGGTGAACATGAGCAACCCCTAC 577
QY      115 GlyAspGlySerGlySerGlyCysTyr-----GlyGly 125
Db      578 GGTGTAACTCAGGCTCTACAGAATACATGTTCCCGGCTCAGCCGGTATTGGCCGGTGGC 637
QY      126 GlyAspGlyGly-----GlyGlySerLeuGlyLeuSerMet 137
Db      638 GCGCGCGTGTAGCAATAGCAGCAACAGCAACAGCACTAGCTCCATAGGGTTATCCATG 697
QY      138 IleLysThrTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsn 157
Db      698 ATAAGACATGTTGAGGAACCAACCACTCAGAAAACAACAATAACAACAACAAT 757
QY      158 AlaAlaLysGly-----LeuSerLeuSerMetAsn--- 167
Db      758 GAAAGTGGTGGCAATAGTAGAAGCAGTGTGCAGCAGACTCTATCACTTTCCATGAGTACT 817
QY      168 -----SerSerThr-----SerCysAspAsnAsnAsnAsp 177
Db      818 GGTTCACAATCAAGCACATCACTACCCCTTCTCACTGCTAGTGGGATAATGGAGAGAGT 877
QY      178 SerAsnAsnAsn-----ValValAlaGlnGlyLysThrIleAspAspSer 192
Db      878 TCTTCTGATAACAAACAACACATACCACCGGTGCACCTTGATACAAACCAACCGGAGCC 937
QY      193 ValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArg 212
Db      938 ATTGAACACTGCACCAGAAAGTCCATTGACACTTTTGACAGAGAACTTCTATCTACCGT 997
QY      213 GlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCys 232
Db      998 GGTGTAAACAAGCATAGGTGGACGGGAGGTATGAGGCTCACCTGTGGGATAATAGTTGT 1057
QY      233 LysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyTyrAspLysGlu 252
Db      1058 AGAAGAGAGGGCAAACTCGCAAGAGGAGGCAAGTTTACTTGGGAGGTTTATGACAAAGA 1117
QY      253 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThrThr 272
Db      1118 GAARAGGCAGCTAGAGCTACGATTTGGCAGCACATAAATACTGGGGAACAACATACGACA 1177
QY      273 ThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLysHisMetThrArg 292
Db      1178 ACAAAATTTTCCAAATTAGCCACTATGAGAAAGAGTTTGGAAAGAAATGAAGCACATGACTAGG 1237
QY      293 GlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIle 312
Db      1238 CAAGAGTACCTTGGCTCATTTGAGAAGGAGAGTAGTGGGTTTCTCGGGGCACTCCATT 1297
QY      313 TyrArgGlyValThrArgHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgVal 332
Db      1298 TATCGAGGTGTACGAGACACCATCAATGGAAGATGCGCAAGCGAGGATTTCGAAGAGTT 1357
QY      333 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluAlaAlaGluAla 352
Db      1358 GGTGGCAACAAGGATCTTACTTGGAACTTTCAGCACCCCAAGAGGAGGACGACAGAGCA 1417
QY      353 TyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsn 372

```

```

Db 1418 TATGATGACGACCAATCAATTCAGAGCACTAGTGTGTTTCAAACTTTGACATGAC 1477
Qy 373 ArgTyrAsnVallyAlaAlaIleuGluSerProSerLeuProIleGlySerAlaAlaIlys 392
Db 1478 AGATATGACGTGAAAGCATACTTGGAGACCACTTTGCCAATTTGGTGGTGTGCAAG 1537
Qy 393 ArgLeuLysGlu-----AlaAsnArg 399
Db 1538 CTTTGAAGCATGACGACAGGTGGAACCTGAGGTGGAGATGTTTCATAGAGCAGATCAA 1597
Qy 400 ProValProSerMetMetMetMetMetMetMetMetMetMetMetMetMetMetMetMet 419
Db 1598 GAAGATCATAGTAGCATCATGAACCTCTCACCTAATCAAGCAATCAATTAACAACATGCA 1657
Qy 420 GlyTyrGlnAsnAlaAlaValGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 439
Db 1658 GCAGGAGGAACAACAGCAGCACTCATCATCACTGCGCAATGCTTGTGCTTCCACCAA 1717
Qy 440 HisGlnGluArgTyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAla 459
Db 1718 CCTCAACCTTGACACCACTACACTACCTTATGACAAAGAAATAT----- 1765
Qy 460 ArgAlaCysPheLysGlnGluAspGln----- 469
Db 1766 -----TGGTGAAGCAAGAACAGACAACTCTGATGCTCTCACTCTTGTCTTATTCA 1819
Qy 470 -----HisPheLeuSerAsnThrGln 477
Db 1820 GATATTTCATCACTACAGCTAGGGAACAAATGGCACACAACTTCTTTCACAAATTC 1879
Qy 478 SerLeu-----MetThrAsnIleAspHisGlnSerSerValSerAspSerValThr 495
Db 1880 GGGTGGACCCCTATGTTAAGCATGGAT-----TCTGCTTCCATGCAATAGTCT 1930
Qy 496 ValCysGlyAsnValVal-----GlyTyrGlyGlyTyrGlnGlyPhe---AlaAlaPro 512
Db 1931 TCATCTAACTCTGTGTTTATGATGTTATGGAGTGGTGGGGCTATAATGATGATTCCT 1990
Qy 513 Val---AsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsn----- 526
Db 1991 ATGGGGACTACTACTCTGTTGTTGTCAAATGATGGTGATCAAAATCCAGAGCAATCAT 2050
Qy 526 ----- 526
Db 2051 GGTGTTGTGATATGACATTAAGCACTGGTTATGAAAGTGTGATGTTCTCAAACT 2110
Qy 527 -----AlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
Db 2111 GATCCTTATCATGCATGCAAGCACTGTTATATCTTACTCAACAGCAACCATCTTCT 2170
Qy 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
Db 2171 -----GTTGATGCAAGTGAAGGCTAGTGCATAT 2197
Qy 561 TyrHisGlyGlyGlyGlyGlyGluValAlaAlaProThr 572
Db 2198 GATCAAGGATCTGCATGCACTACTTGGTTCCTCAACT 2233

```

RESULT 3

```

US-10-183-687-361
; Sequence 361, Application US/10183687
; Publication No. US20030204870A1

```

GENERAL INFORMATION:

```

; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cannon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted

```

```

; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BE1458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 361
; LENGTH: 1990
; TYPE: DNA
; ORGANISM: Glycine max
US-10-183-687-361

```

```

Alignment Scores:
Pred. No.: 2,65e-120 Length: 1990
Score: 1119.50 Matches: 275
Percent Similarity: 55.70% Conservative: 77
Best Local Similarity: 43.51% Mismatches: 147
Query Match: 35.89% Indels: 133
DB: 13 Gaps: 23

```

```
US-09-980-364-2 (1-579) x US-10-183-687-361 (1-1990)
```

```

Qy 1 MetAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLys 20
Db 126 ATGAACAACAACACTGGCTTTTCGTTCCCTCTTCT----- 158
Qy 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
Db 158 ----- 158
Qy 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
Db 159 CCTACTCATCTCTCTTACCACTCATCTCTCAAGCACTCAATATCATCAATTTTCC 218
Qy 61 ValValValAlaPheThrArgAspAsnAsnSerHisSerArgAspTyrAsp---Ile 79
Db 219 CTGGTGTAGTGAACAGAGAACATGGATAACCTTTCCAAAATCATGATTGAATCTGATT 278
Qy 80 AsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsn 99
Db 279 AACACCCATAGTAGCAACGAAAT-----CCAAAAGTGGCTCAT 317
Qy 100 PheLeuGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 119
Db 318 TTTCATAGCA-----GTGAGCAAGTCTGAAAAATCAGTCAGACCTTTGAGACC 362
Qy 120 SerGlyCysTyrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 139
Db 363 TTAACAGAAATTCATTCAAATGATTCAGATATCTGTTCCAAAACAACAGCTCTGGTGCCT 422
Qy 140 ThrTrpLeuArgAsnGlnProValAspAsnValAspAsn-----GlnGluAsnGlyAsn 157
Db 423 -----ATGCAAAAACCCCTGTGTGGACACACCTAGCAATGAGTATCAAGAAAATGCTAAT 476
Qy 158 Ala---AlaLysGlyLeuSerLeuSerMetAsnSer-----SerThrSerCysAspAsn 174
Db 477 AGTAATTTGCAATCATTCACATTCATTCATGGGAAGTGTGAAGGATTCACATGTGAAC 536
Qy 175 AsnAsnAspSerAsnAsnAsnValValAlaGlnGlyLysThrIleAspAspSerValGlu 194
Db 537 AGTGTGAAAAATAGCAAAACACT-----ACTGTTGAA 569
Qy 195 AlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyVal 214
Db 570 GTTGACACCTAGAAACCTTTGGATACATTCGGGAGAGAACATCCATATATCGTGGAGTA 629

```



```
QY 215 ThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysIysArg 234
Db 630 ACTCGACATAGATGGACTGCAAGGTATGAAGCTCATCTTTGGGATAATAGCTGTAGAAGG 689
QY 235 GluGlyGlnThrArgIysGlyArgGlnValTyrLeuGlyGlyTyrAspIysGlnGluIys 254
Db 690 GAAGGCCAATCAAGAAAGAGCGCCCAAGTTTATTTGGGTGGATATGATTAAGAGAGAAA 749
QY 255 AlaAlaArgAlaTyrAspLeuAlaAlaLeuIysTyrTrpGlyThrThrThrThrAsn 274
Db 750 GCAGCTAGAGCTTATGATTAGCTGCACTGAAGTACTGGGGGACATCCACCCTACCAAC 809
QY 275 PheProMetSerGluTyrGlnIysGluValGluGluMetIysHisMetThrArgGlnGlu 294
Db 810 TTTCCAAATAGCAACTATGAGAGGAATTTGGATGAATGAACACATGACGAGCAAGAA 869
QY 295 TyrValAlaSerLeuArgIysSerSerGlyPheSerArgGlyValaSerIleTyrArg 314
Db 870 TTTGTTGCCGCCATTAGAGGAAAGACAGTGGTTTCTCCAGGGGTCATCAATGATCGT 929
QY 315 GlyValThrArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGly 334
Db 930 GGAGTTTACAAGGCATCACCAACACGGAAGATGCAAGCAAGGATTGGCAGAGTTGCAGA 989
QY 335 AsnIysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaTyrAsp 354
Db 990 AACAAAGATCTTTACTTTGGAACTTTTCAGTACTGAGGAAGAGGCTGCAGAGCATACGAC 1049
QY 355 IleAlaAlaIleIysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyr 374
Db 1050 ATAGCAGCGATAAAGTTACAGGCTCTCAACGCTGTCAACAACTTTGACATGAGCCGCTAC 1109
QY 375 AsnValIysAlaIleLeuGluSerProSerLeuProIle---GlySerAlaAlaIysArg 393
Db 1110 GAGCTGAAGCCATCTTTGAAGCAACACACTCTCCCAATAGGAGGAGCGCTGCAAGCGT 1169
QY 394 LeuIysGluAla-----AsnArgProValProSerMetMetIle----- 407
Db 1170 CTGAAGAAGCTCAAGCTCTAGAATCTTCGAGAAAACGCGAAGAGATGATGCTACTAGGC 1229
QY 408 SerAsnAsnValSerGluSerGluAsnSerAlaSerGlyTrpGlnAsnAlaAla----- 425
Db 1230 TCATCTTCCAGCTTCCAATACGGAACCTCAGCAAGCTCTTCTAGGCTTCACGCTTACCCT 1289
QY 426 ---ValGlnHisHis-----GlnGlyVal 432
Db 1290 CTAATGAGCAGCACCCACAGTTCGAGCAACCTCACTCTGCTAACTCTTCAAAACCAC 1349
QY 433 AspLeuSerLeuLeuHis---GlnHisGlnGlu-----ArgTyrAsnGlyTyrTyr 448
Db 1350 GACATAAGTTCTTCTCCTCTCTCTCACCAGCAAGACCTTTTGCATCATCATCGGCTTACATC 1409
QY 449 TyrAsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheIys---GlnGluAsp 467
Db 1410 CAAACGCGAGTTCAGTTGCGACGAGTGCGGCTCTTCTTATAGCTTTTCAGATAAT 1469
QY 468 AspGlnHisHis-----PheLeuSerAsnThrGlnSerLeu-----MetThrAsn 482
Db 1470 GCTCAGTTCTACAATAGTTTACCTTCAGAACCCCTGCAITGCTTCAGGGAATGATGAAC 1529
QY 483 IleAspHisGlnSerSerValSer-----AspAspSerValThrValCys 497
Db 1530 ATGGGGTCTTCTTCTTCTCTCTCTGTTGAGAAATTAATATAGTAACAATAAT 1589
QY 498 GlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAla 517
Db 1590 AATAATGTTGGTGGTTTGTGGGA---AGTGGGTTTGGTATGGCTTCGAATGCAACGGCG 1646
QY 518 -----TyrAlaAlaSerGluPheAspTyrAsn 526
Db 1647 GGGAAACAGCGTGGGACAGCGAGGAGTTAGGGCTGGTGGAGGTGGACTATGACATGCCG 1706
QY 527 AlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGlnGlnSerProGlyGlyAsp 546
```

```
Db 1707 GCTGGAGGTTACGGTGGCTGGTGGCGCGGACTCCATCGACAGCTCAATGTGGG--- 1763
QY 547 PheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGly 566
Db 1763 ----- 1763
QY 567 GlyGluValAlaProThrPheThrValTrpAsnAsp 578
Db 1764 -----GTGTTCACAATGTGAATGAT 1784

RESULT 4
US-10-424-599-91337
; Sequence 91337, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 91337
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53488C.1
US-10-424-599-91337
```

```
Alignment Scores:
Pred. No.: 2,628-119 Length: 2169
Score: 1111.50 Matches: 274
Percent Similarity: 54.01% Conservative: 76
Best Local Similarity: 42.28% Mismatches: 142
Query Match: 35.64% Indels: 157
DB: 13 Gaps: 20
```

```
US-09-980-364-2 (1-579) X US-10-424-599-91337 (1-2169)
QY 1 MetAsnAsnAsnTrpIleuGlyPheSerLeuSerProTyrGluGlnAsnHisArgIys 20
Db 133 ATGAACAACAACCTGGCTTTCGTTCCCTCTTCT----- 165
QY 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
Db 165 ----- 165
QY 41 ProThrAlaAlaAspAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
Db 166 CCTACTCATCTTCTTCTTACCAGCTCATGATCTTCAAGCAACTCAATATCATCAATTTTCC 225
QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAsp---Ile 79
Db 226 CTGGGTTAGTGAACGAGAACATGGATACCCCTTCCAAAATCATGATTTGGATCTGAT 285
QY 80 AsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProIysLeuGluAsn 99
Db 286 AACACCCATAGTAGCAACGAAAT-----CCAAAAGTGGCTGAT 324
QY 100 PheLeuGlyArgThrThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
Db 325 TTTCTAGGA-----GTGAGCAAGTCTGAAATCAGTCAGACCTTCGAGCC 369
QY 120 SerGlyCysTyrGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleIys 139
Db 370 TTAACGAATTCATTCATCAATGATTCAGATTATCTGTTCAACAACACAGCTGTGGCT 429
QY 140 ThrTrpLeuArgAsnGlnProValAspAsnValAspAsn-----GlnGluAsnGlyAsn 157
```



```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 1738
; LENGTH: 2088
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G1793
US-10-374-780A-1738

```

```

Alignment Scores:
Pred. No.: 3,23e-119 Length: 2088
Score: 1110.50 Matches: 274
Percent Similarity: 53.86% Conservative: 75
Best Local Similarity: 42.28% Mismatches: 139
Query Match: 35.60% Indels: 161
DB: 16 Gaps: 20

```

US-09-980-364-2 (1-579) x US-10-374-780A-1738 (1-2088)

```

QY 1 MetAsnAsnThrLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
Db 133 ATGAACACCAACTGGCTTCGTTCCCTCTTCT- 165
QY 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
Db 165 ----- 165
QY 41 ProThrAlaAlaSerAspGluSerAlaIleGlnThrSerPheProSerProPheGly 60
Db 166 CTAATCACTTCCTTACCAGCTCATGCTTCACGCACTCAATCATCAATTTCC 225
QY 61 ValValValAspAlaPheThrArgAspAsnSerHisSerArgAspTrpAsp 79
Db 226 CTTGGGTTAGTGAACGAGAACATGATTAACCTTTCCAAATCATGATTGGAATCTGATT 285
QY 80 AsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsn 99
Db 286 AACACCCATAGTACGACGAAAT- 324
QY 100 PheLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
Db 325 TTCTTAGGA-----GTGACCAAGTCTGAAATCAGTACAGACCTTGCAGCC 369
QY 120 SerGlyCysTyrGlyGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLys 139
Db 370 TTAACGGAATTCATTCATCAATGATTCAGATATCTGTTCACAAACACAGCTCTGGTGCCT 429
QY 140 ThrTrpLeuArgAsnGlnProValAspAsnValAspAsn-----GlnGluAsnGlyAsn 157
Db 430 -----ATGCAAAACCCCTGTTGGACACACCTAGCAATCAGTATCAAGAAATGCTAAT 483
QY 158 Ala---AlaLysGlyLeuSerLeuSerMetAsnSer-----SerThrSerCysAspAsn 174
Db 484 AGTAATTTGCAATCATTTGACATTCATCCATGGGAGTGGTAAGGATTCACATGTGAAC 543
QY 175 AsnAsnAspSerAsnAsnValValAlaGlnGlyLysThrIleAspAspSerValGlu 194
Db 544 AGTGGTGAATATACCAACAACT- 576
QY 195 AlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyVal 214
Db 577 GTTGACCTTCAAGAACCTTTGGATACATTCGGGACAGAACATCCATATATCTGTGGAGTA 636
QY 215 ThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArg 234
Db 637 ACTGCACATAGTACTGGAGGTATGAAGCTCATCTTTGGGATATATAGCTGTAGAGG 696
QY 235 GluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLys 254
Db 697 GAAGGCCAATCAAGAAAGGACGCCAAGTTTATTTGGGTGGATATGATAAAGAGAGAAA 756
QY 255 AlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsn 274

```

```

Db 757 GCAGCTAGGCGTTATGATTTAGCTGCACTGAAGTACTGGGGACATCCACCACTACCAAC 816
QY 275 PheProMetSerGluTyrGluLysGluValGluMetLysHisMetThrArgGlnGlu 294
Db 817 TTTCCAAATTTAGTAATGAGAGGAATTCGATGAATGAACACATGACGCGACAGAA 876
QY 295 TyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArg 314
Db 877 TTTGTTGCTGCCATTAGAGGAAAGCAGTGGTTTCTCCAGGGGTGCATCAATGTATCTGT 936
QY 315 GlyValThrArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGly 334
Db 937 GGAGTTACAGGCAATCCACACACGGAAGATGGACACGAGAATTTGCAGAGTTGCAGGA 996
QY 335 AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAsp 354
Db 997 AACAAAGATCTTTACTTTGGGAACCTTCAGTACTCAAGAAGAGGCTGTGAAGCATACGAC 1056
QY 355 IleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyr 374
Db 1057 ATAGCTGCGATAAAGTTTCAGAGGTTCTCAACGCTGTCAAAACTTTTGACATGAGCGGTAC 1116
QY 375 AsnValLysAlaIleLeuGluSerProSerLeuProIle---GlySerAlaAlaLysArg 393
Db 1117 GACGTGAAGGCCATCTTTGAAGCAACACTCTCCCAATAGGAGGAGCTGCAAGCGGT 1176
QY 394 LeuLysGluAla-----AsnArgProValProSerMetMetMetIleSer 408
Db 1177 CTGAAGAAAGCTCAAGCTCTAGAATCTTCGAGAAGCGCGAAGATGATTCACACTAGGA 1236
QY 409 AsnAsnValSerGluSerGluAsnSerAlaSerGlyTrpGlnAsnAlaAla----- 425
Db 1237 TCATCCACATTCCTCAATATGGAAACCAACAGCTCTAATTTCTAGGCTACATGCTTACCCTCTA 1296
QY 426 ValGlnHisHis-----GlnGlyValAsp 433
Db 1297 ATGCAGCACCAACCACTGAAACCACTCAACCTCTGCTAATCTTTCGAAACCACTGAT 1356
QY 434 LeuSerLeu-----LeuHisGln----- 439
Db 1357 ATCATGTTCTCACTTCTCTCACCAGCAAGACCTTTTGCATCAGGTTTACATCCAAACGCG 1416
QY 440 -----HisGlnGluArg----- 443
Db 1417 CTTCACTGGTGCACAGCAGCAGAGTGGTGGTTCTTCTTCTTATAGCTTTTCAGAAATAAT 1476
QY 444 ---TyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCys 462
Db 1477 ATAAATAATGCTCAGTCTCTATAATGGTTATATCTCAGAACCCACCTGCTGCT--- 1533
QY 463 PheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMetThrAsn 482
Db 1534 -----CAGGGAATGATTAACTGGGTCTTCATCTTCTTCATCTGTG 1575
QY 483 IleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnValValGly 502
Db 1576 TTGGAGAAATAATAAGTACCAATAATAATGTGG- TGGGTTTGTGGGAAGTGGGTTGG 1634
QY 503 TyrGlyGlyTyrGlnGlyPheAlaProValAsnCysAspAlaTyrAlaAlaSerGlu 522
Db 1635 TATGGC----- 1640
QY 523 PheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGlnGlnSer 542
Db 1641 TTCTAATGCAACGTCGGGGAACAC----- 1664
QY 543 ProGlyGlyAspPheProAlaAlaMetThrAsn-AsnValGlySerAsnMetTyrTyrHis 562
Db 1665 ---GGTGGGAGCGGGAGGAGCTAGGGCTGTGTGAGGTGGACTATGACATGCCACTGG 1721
QY 562 sGlyGluGlyGlyGluValAlaPro----- 571
Db 1722 TGGTTACGGTGGATGTCGGCGCGCGCGGAGTCCATGCAGACGTCGAATAGTGG 1781

```

QY 572 -ThrPheThrValTrpAsnAsp 578
 DB 1782 GGTGTTCCACNATGTGGATGAC 1803

RESULT 6
 US-10-183-687-353
 ; Sequence 353, Application US/10183687
 ; Publication No. US20030204870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Allen, William B.
 ; APPLICANT: Cahoon, Rebecca
 ; APPLICANT: Epeibaum, Sabine
 ; APPLICANT: Famodu, Omolayo O.
 ; APPLICANT: Harvelli, Leslie T.
 ; APPLICANT: Jones, Todd
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Klein, Ted
 ; APPLICANT: Li, Changjiang
 ; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczynski, Mitchell C.
 ; TITLE OF INVENTION: Alteration of Oil Traits in Plants
 ; FILE REFERENCE: BBI458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; PRIOR FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 353
 ; LENGTH: 1824
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-183-687-353

Alignment Scores:
 Pred. No.: 3,38-115 Length: 1824
 Score: 1075.50 Matches: 240
 Percent Similarity: 59.62% Conservative: 45
 Best Local Similarity: 50.21% Mismatches: 98
 Query Match: 34.48% Indels: 95
 DB: 13 Gaps: 12

US-09-980-364-2 (1-579) x US-10-183-687-353 (1-1824)

QY 167 AsnSerSerThrSerCysAspAsnAsnAspSerAsnAsnValAlaGlnGly 186
 DB 12 AATGAGAGAGTCTCTCTGATAACAAACCA-----AACACCTCGGCTGCATTGAT 65
 QY 187 LysThrIleAspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGln 206
 DB 66 TCCACCAACCGGAGCATGAACTGCACCCAGAAAGTCCATTGACACTTTTGACACAG 125
 QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226
 DB 126 AGAAGCTTCTATCTACCGTGTGTAAACAGCATAGTGGACGGGGAGGTACGAGGCTCAC 185
 QY 227 LeuTrpAspAsnSerCysLysArgGluGlnThrArgLysGlyArgGlnValTyrLeu 246
 DB 186 CTGGGGATATAGTTGTAGAGAGAGCGACAGCTCGCAAGGAAAGGCAAGTTTACTTG 245
 QY 247 GlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyr 266
 DB 246 GGTGGTATGATAAGAGAAAGGACGCTAGAGCCTACGATTGGCAGCCTAAATATAC 305
 QY 267 TrpGlyThrThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
 DB 306 TGGGGAACCAACCAACAAATTTTCCAAATTAGCCACTATGAGAAAGAGTTTGAAGAA 365
 QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPhe 306

DB 366 ATGAAGCACATGCTAGGCAAGAGTAGTGGTCACTTGGAGAGAGAGTAGTGGT 425
 QY 307 SerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGln 326
 DB 426 TCTCGCGGTGCATCCATTTATCGAGGAGTGAGGAGACACCAACAGAGTGGCAA 485
 QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
 DB 486 GCGAGGATTGGAGAGTCTGCGCAACAGAGTCTTTACTTGGAACTTTTAGCAACCAA 545
 QY 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
 DB 546 GAAGAGCGAGCGAAGCATATGATGATGAGCAGCAATCAAAATTCGAGGAGTAAAGTCTGT 605
 QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluLysProSerLeuPro 386
 DB 606 ACAAACTTTGACATGAGCAGATATGACGTGAAAGCATACTTGAGAGCACCACCTTTGCCA 665
 QY 387 IleGlySerAlaAlaLysArgLeuLysGlu----- 396
 DB 666 ATAGTGTGTGTCGCAAGCGTTTGAAGGATATGAGGAGGAGTGAACCTGAGTGTGATAAT 725
 QY 397 -----AlaAsnArgProValProSerMetMetMetIleSerAsnValSerGlu 413
 DB 726 GGTATAGAGCAGATCAAGTAGATCATATGATCATGAGTCTTCACATACTCAAGGA 785
 QY 414 SerGluAsnSerAlaSerGlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAsp 433
 DB 786 ATCAATAACAACCTATGAGGAGGGGGAACAGCACT-----CATCATACTGGCACAAT 839
 QY 434 LeuSerLeuLeuHisGlnHisGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsn 453
 DB 840 GCTCATGCAATTCACCAACCTCAACCTTGCACCACTGACCTACCTTATGGACAAAGA 899
 QY 454 LeuSerSerGluSerAlaAlaCysPheLysGlnGluAspAspGln----- 469
 DB 900 ATTAAT-----TGGTGAAGCAAGCAACCAACCAAGCAACTCTGATGCCCT 944
 QY 470 -----HisHis 471
 DB 945 CACTCTTGTCTTATTATTCAGATATTCACTCACTTCAGCTAGGGAACAATGGAACATAAC 1004
 QY 472 PheLeuSerAsnThrGlnSerLeu-----MetThrAsnIleAspHisGlnSerVal 489
 DB 1005 TTCTTTCACAAAATTCAGGGTTCACCTCTGTTGAGCATGGAT-----TCTGCT 1055
 QY 490 SerAspAspSerValThrValCysGlyAsnValVal-----GlyTyrGlyTyrGln 507
 DB 1056 TCCATTGACAATAGTCTCTCTTAACCTGGTGTATTATGATGGTATGAGGTTGGTGG 1115
 QY 508 GlyPhe---AlaAlaProVal---AsnCysAspAlaTyrAlaAlaSerGluPheAspTyr 525
 DB 1116 GCGTCAATATGATGCTTATGAGAACTACTACTCTGTGTGCAAGTATGATGGTGAACAA 1175
 QY 526 Asn----- 526
 DB 1176 AATCCAAGAACCAATCATGGTTTGGTGATAATGAGATAAAGCACCTGGTTATGAAAGT 1235
 QY 527 -----AlaArgAsnHisTyrTyrPheAla 534
 DB 1236 GTGTATGGCTCTGCAACTGATCTTATCATGCATGCAAGAACTGTATTATCTTACT 1285
 QY 535 GlnGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsn 554
 DB 1286 CAACAGCAA-----TCATCTTCTGTGTGATACA 1322
 QY 555 ValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaProThr 572
 DB 1323 GTGAAGGCTAGTGCATATGATCAAGGGTCTGCACTGCACTACTTGGTTCACAACT 1376

RESULT 7
 US-10-225-066A-365

```

; Sequence 365, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBEL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: FILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 365
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-365

Alignment Scores:
Pred. No.:      6,26e+112      Length:      1818
Score:          1047,50       Matches:     262
Percent Similarity: 46.96%    Mismatches:   70
Best Local Similarity: 37.06%  MisMatches:  110
Query Watch:     33.58%      Indels:       265
DB:              13           Gaps:         23

US-09-980-364-2 (1-579) x US-10-225-066A-365 (1-1818)

Qy      2 AsnAsnAsnTrpLeuGlyPhe-----SerLeuSerProTyrglu 14
      : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db      65 TCTAACAACTGGCTGGGTTCCTTTTACCGACAACACTCTCTTGCCCTCCTCATGAA 124
      : : : | | | | | | | | | | | | | | | | | | | | | | | |

Qy      15 GluAsnHisArgLysAspValTyrsSerThrThrThrValValAspValala 34
      : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db      125 TACAAC-----130

Qy      35 GlyGluTyrcysTyrsAspProThralalaserAspGluSerSerAlalleGlnThrSer 54
      : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db      130 -----130

Qy      55 PheProcSerProPheGlyValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
      : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db      131 -----CTTGGCTTGGTCAGCACCATATG-----GACAACCCTTTTCAACA 172
      : : : | | | | | | | | | | | | | | | | | | | | | | | |

Qy      75 ArgAspTrpAspileAsnGlycysAlaCysAsrAsnilleHisAsnAspGluGlnAspGly 94
      : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db      173 CAAGATGGAAATG-----ATCAATCCACACCGTGGAGGAGGATGAA 217
      : : : | | | | | | | | | | | | | | | | | | | | | | | |

Qy      95 -----ProlysleuGluasnPhelleuGlyArgThrThrThrIleTyrsenThr 110
      : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db      218 GGAGGAGAGGTTCCAAAAGTGCAGCATATTTTCGGGTGTGAGCAACCCGCCGACGAACCAA 277
      : : : | | | | | | | | | | | | | | | | | | | | | | | |

Ov      111 AsnGluAsnValglyVaspolvsSerglyscysTyrglygylAspglyglygly 130

```

Db 1310 COATTAGACGCTTTCTATCTCTTCAGAACATGACATCTCTCAATTAACAACAACAAT 1369
 QY 420 GlyTyrGlnAsnAlaValGlnHisGlnGlyValAspLeuSerLeu---LeuHis 438
 Db 1370 GCTCAGATCTCTCTCTTTAATCACCATAGCTATATCCAGACAACTTCACTCCAC 1429
 QY 439 GlnHis-----GlnGluArgTyrAsn 445
 Db 1430 CAACAGACCAACAATTACTTTCAGCAACAGTCGAGCCAGAACTCTCAGCAGCTTCAAT 1489
 QY 446 GlyTyrTyrTyr----- 449
 Db 1490 GGTATCTTCATAGCAATCCGCTCTGCTTCATGACTTGCTCTACCTCTATCGTTGAC 1549
 QY 450 -----AsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheLysGln 465
 Db 1550 AACAAATAATAACAATGAGGCTCTAGTGGAGCTACAACACTGCGCA----- 1597
 QY 466 GluAspAspGlnHisPheLeuSerAsn----- 475
 Db 1598 -----TTCTTGGGAACCAACGGTATGTATGGTCCAGCTCGACT 1639
 QY 476 -----ThrGlnSerLeuMetThr---AsnIleAspHisGlnSerSerValSerAsp 491
 Db 1640 GTTGGATCGACCGAGGAGTTCCACCGTTAAACAGATTACGATATGCCTTCCAGTGAT 1699
 QY 492 AspSerValThrValCysGlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAla 511
 Db 1700 -----GGAACCGAGGGTATAGTGTGG----- 1723
 QY 512 ProValAsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArgAsnHisTyr 531
 Db 1723 ----- 1723
 QY 532 TyrPheAlaGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMet 551
 Db 1724 -----ACCAAGTGTCTGTTCCAGGGGTCAAAACCT----- 1753
 QY 552 ThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyValAlaPro 571
 Db 1754 -----GGTGGT----- 1759
 QY 572 ThrPheThrValTrpAsnAsp 578
 Db 1760 GTTTTCACTATGTGAATGAG 1780

RESULT 8

US-10-225-066A-855
 ; Sequence 855, Application US/10225066A
 ; Publication No. US20030226173A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: RIECHMANN, Jose Luis
 ; APPLICANT: ADAM, Luc J
 ; APPLICANT: DUBELL, Arnold T
 ; APPLICANT: HEARD, Jacqueline E
 ; APPLICANT: PILGRIM, Marsha L
 ; APPLICANT: JIANG, Cai-Zhong
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: CREELMAN, Robert A
 ; APPLICANT: PINEDA, Omaira
 ; APPLICANT: YU, Guo-Liang
 ; APPLICANT: BROUN, Pierre E
 ; TITLE OF INVENTION: Field-Related Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: MB10036-2 US
 ; CURRENT APPLICATION NUMBER: US/10/225,066A
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 09/837,444
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/310,847
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-12-05
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 1122
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 855
 ; LENGTH: 1818
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-225-066A-855

Alignment Scores:
 Pred. No.: 6,26e-112 Length: 1818
 Score: 1047.50 Matches: 262
 Percent Similarity: 46.96% Conservative: 70
 Best Local Similarity: 37.06% Mismatches: 110
 Query Match: 33.58% Indels: 265
 DB: 13 Gaps: 23

US-09-980-364-2 (1-579) x US-10-225-066A-855 (1-1818)

QY 2 AsnAsnAsnTrpLeuGlyPhe-----SerLeuSerProTyrGlu 14
 Db 65 TCTAACACTGGCTTGGCTTCTCTTTCACGGAACAACCTCTTTCCTCTCATGAA 124
 QY 15 GlnAsnHisHisArgLysAspValTyrSerSerThrThrValValAspValAla 34
 Db 125 TACAAC----- 130
 QY 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleClnThrSer 54
 Db 130 ----- 130
 QY 55 PheProSerProPheGlyValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
 Db 131 -----CTTGGCTTGGTTCAGCACCATATG-----GACAACCTTTTCAACA 172
 QY 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94
 Db 173 CAAGAGTGAATATG-----ATCAATCCACACACGGTGGAGGAGGAGATGAA 217
 QY 95 -----ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThr 110
 Db 218 GGAGGAGAGGTTCCAAAAGTGGCGGATTTCTCGTGTGAGCAAAACCGGAGCAAAACCA 277
 QY 111 AsnGluAsnValGlyAspGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGly 130
 Db 278 TCCAAACCACCTAGTAGCTTACAACAGCTCAGACTACTACTTCCATACCAAT----- 328
 QY 131 GlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProVal----- 147
 Db 329 -----AGCTTGCTAGCTGCTCCAAATCAAAACGATGCTGTTGTAGCAGCT 373
 QY 148 -----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
 Db 374 TGTGACTCCAATCTCTTAACAACAGTAGTATCATGAGCTTCAAGAGAGTCTCACAAT 433
 QY 159 AlalysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
 Db 434 CTACAGTCACCTACTTGTTCATGGGACCAACC-----GCT 469
 QY 179 AsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu--- 194
 Db 470 GGTAATATGTTGTAGACAAAGCTTCACCTCCAGACCACCGGGGATAACGCTAGCGGT 529
 QY 195 -----AlaThrProLysLysThrIleGluSerPheGlyGln 206
 Db 530 GGAGCAGTACCGGTGTTGAGACGCGCCACGCAAGACGTCATTTGGACACTTTGGGCAA 589
 QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226

Db 65 TCTAACAACTGGCTGGCTTCTCTTCCACCGAAACAACCTCTCTTCTGGCTCTCATGAA 124
 QY 15 GlnAsnHisHisArgIysAspValTyrSerSerThrThrThrValValAspValAla 34
 Db 125 TACAAC----- 130
 QY 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54
 Db 130----- 130
 QY 55 PheProSerProPheGlyValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
 Db 131-----CTTGGCTTGTTCAGCACCATATG-----GACACCTTTTCAACAA 172
 QY 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94
 Db 173 CAAGAGTGAATATG-----ATCAATCCACACGGTGGAGGAGGAGATCAA 217
 QY 95-----ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThr 110
 Db 218 GGAGGAGAGGTTCCAAAAGTGGCCGATTTCTCGTGTGACAAACCGGACGAAACCAA 277
 QY 111 AsnGluAsnValGlyAspGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGly 130
 Db 278 TCCAAACCACCTAGTAGCTTACAACGACTCAGACTACTACTTCCATACCAAT----- 328
 QY 131 GlySerLeuGlyLeuSerMetIleIysThrTrpLeuArgAsnGlnProVal----- 147
 Db 329-----AGTTTATGCTAGCTGCTCAACGCTCAACATCAACGATGCTGTGTAGCAGCT 373
 QY 148-----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
 Db 374 TGTGACTCCAATACTCTCAACAGTAGCTATCATGAGCTTCAAGAGAGTGCTCACAAT 433
 QY 159 AlaySerGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
 Db 434 CTACAGTCACCTTACTTTGTCCATGGGACCACC-----GCT 469
 QY 179 AsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu--- 194
 Db 470 GGTAAATATGTTAGACAAGCTTCCACCTCGAGACACCGGGGATAACCGTAGCGGT 529
 QY 195-----AlaThrProLysLysThrIleGluSerPheGlyGln 206
 Db 530 GGAGCACTAGCGTGTGTGACGCGCCACGCAAGACGTCATGGACACTTTCCGACAA 589
 QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226
 Db 590 CGAACCTCGATCTATCTGTGTGTGCACAGACATCGATGGACTGCTCGATATGAGGCTCAT 649
 QY 227 LeuTrpAspAsnSerCysLysArgGlyGlnThrArgLysGlyArgGlnValTyrLeu 246
 Db 650 CTATGGCAATAATAGTTAGAGGAAGGAGCGAGTCTAGGAAAGAGACAAAGTTTACTTG 709
 QY 247 GlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaIleLeuLysTyr 266
 Db 710 GGTGGATATGACAAAGAGATGAAGCAGCAAGATCATATGATCTAGTCGACATTAAGTAC 769
 QY 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
 Db 770 TGGGGTCTCTCAACTACTACTATTTCCCATTAACAACCTACGAGAAAGAGTAGAGGA 829
 QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPhe 306
 Db 830 ATAAGCACATGACGACAGCAAGAGTTCGTGGCTGCCATTAGAGGAAAGATAGTGATTT 889
 QY 307 SerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGlyArgTrpGln 326
 Db 890 TCGAGAGGCCCTTCGATGATCGAGAGATACAGGCATCACCAACATCGAAGATGGCAA 949
 QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
 Db 950 GCAAGGATCGCCGAGTGGCCGGAACAAAGACCTCTACTTGGGAACCTTTTACGACTGAG 1009

QY 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
 Db 1010 GAAGAAGCAGCAGAAAGCTTACGATATAGCTGCAATAAAGTTTAGAGGACTTAATGAGTG 1069
 QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
 Db 1070 ACCAACTTCGAGATCAACCGGTACGAGTGAAGCAATTTCTAGAGAGTAGCACTCTTCC 1129
 QY 387 Ile---GlySerAlaAlaLysArgLeuLysGluAla----- 397
 Db 1130 ATCGGAGGAGGCGCAGCTAAACGGCTCAAAGAAGTCAAGCTCTTTCAGTCTTTCAGAGGAA 1189
 QY 397----- 397
 Db 1190 CGCGAGCGGAGATGATAGCCCTTGGTTCAAGTTTCCAGTACGGTGGTGGCTCGAGCACA 1249
 QY 398-----AsnArg 399
 Db 1250 GGCTCTGGCTCCACCTCATCAAGACTTTCAGCTTCAACCTTACCTCTTAAGCATTTCAACA 1309
 QY 400 ProValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSer 419
 Db 1310 CCATTAGAGCCTTTTCTATCTCTTCAGAACATGACATCTCTCATTAACAACAACAAAT 1369
 QY 420 GlyTrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeu---LeuHis 438
 Db 1370 GCTCAGCATTCCTCTCTTTTAAATCACCATAGTATATCCAGACACAACCTTCATCTCCAC 1429
 QY 439 GlnHis-----GlnGluArgTyrAsn 445
 Db 1430 CAACAGCAACAACATTAATTCAGCAACAGTCGAGCCAGAACTCTCAGCAGCTCTACAAT 1489
 QY 446 GlyTyrTyrTyr----- 449
 Db 1490 GCGTATCTTCATAGCAATCCGGCTCTGCTTCATGGACTGTCTCTACCTCTATCGTTGAC 1549
 QY 450-----AsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheLysGln 465
 Db 1550 AACAAATATAACAATGGAGGCTCTAGTGGAGCTACACACTGCAGCA----- 1597
 QY 466 GluAspAspGlnHisHisPheLeuSerAsn----- 475
 Db 1598-----TTTCTTGGGAACCAACCGGTATGGTATGGTCCGCTCGACT 1639
 QY 476-----ThrGlnSerLeuMetThr---AsnIleAspHisGlnSerSerValSerAsp 491
 Db 1640 GTTGATCGACCGAGGAGTTCACACCGTTAAACAGATACGATATGCCCTTCCAGTGAT 1699
 QY 492 AspSerValThrValCysGlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAla 511
 Db 1700-----GGAACCGGAGGTATAGTGTGG----- 1723
 QY 512 ProValAsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArgAsnHisTyr 531
 Db 1723----- 1723
 QY 532 TyrPheAlaGlnGlnGlnThrGlnSerProGlyGlyAspPheProAlaAlaMet 551
 Db 1724-----ACCAGTGAGTCTCTTCCAGGGGTCAACCCCT----- 1753
 QY 552 ThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaPro 571
 Db 1754-----GGTGGT----- 1759
 QY 572 ThrPheThrValTrpAsnAsp 578
 Db 1760 GTTTTCACTATGTGGAATGAG 1780

RESULT 10

US-10-374-780A-333

; Sequence 333, Application US/10374780A

; Publication No. US20040019927A1

GENERAL INFORMATION:
 APPLICANT: Sherman, Bradley K
 APPLICANT: Riechmann, Jose Luis
 APPLICANT: Jiang, Cai-Zhong
 APPLICANT: Heard, Jacqueline E
 APPLICANT: Heake, Volker
 APPLICANT: Creelman, Robert A
 APPLICANT: Ratcliffe, Oliver
 APPLICANT: Adam, Luc J
 APPLICANT: Reuber, T. Lynne
 APPLICANT: Keddle, James
 APPLICANT: Bryon, Pierre E
 APPLICANT: Pilgrim, Marsha L
 APPLICANT: Dubell III, Arnold T
 APPLICANT: Pineda, Omaira
 APPLICANT: Yu, Guo-Liang
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
 FILE REFERENCE: MBI-0047 CIP
 CURRENT APPLICATION NUMBER: US/10/374,780A
 CURRENT FILING DATE: 2003-02-25
 PRIOR APPLICATION NUMBER: 09/837,944
 PRIOR FILING DATE: 2001-04-18
 PRIOR APPLICATION NUMBER: 60/310,847
 PRIOR FILING DATE: 2001-08-09
 PRIOR APPLICATION NUMBER: 09/934,455
 PRIOR FILING DATE: 2001-08-22
 PRIOR APPLICATION NUMBER: 60/336,049
 PRIOR FILING DATE: 2001-11-19
 PRIOR APPLICATION NUMBER: 60/338,692
 PRIOR FILING DATE: 2001-12-11
 PRIOR APPLICATION NUMBER: 10/171,468
 PRIOR FILING DATE: 2002-06-14
 PRIOR APPLICATION NUMBER: 10/225,066
 PRIOR FILING DATE: 2002-08-09
 PRIOR APPLICATION NUMBER: 10/225,067
 PRIOR FILING DATE: 2002-08-09
 PRIOR APPLICATION NUMBER: 10/225,068
 PRIOR FILING DATE: 2002-08-09
 NUMBER OF SEQ ID NOS: 2906
 SOFTWARE: PatentIn version 3.2

SEQ ID NO 333
 LENGTH: 1818
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana

FEATURE:
 OTHER INFORMATION: G1793

US-10-374-780A-333

Alignment Scores:

Pred. No.: 6.26e-112 Length: 1818

Score: 1047.50 Matches: 262

Percent Similarity: 46.96% Conservative: 70

Best Local Similarity: 37.06% Mismatches: 110

Query Match: 33.58% Indels: 265

DB: 16 Gaps: 23

US-09-980-364-2 (1-579) x US-10-374-780A-333 (1-1818)

QY 2 AsnAsnAsnTrpLeuGlyPhe-----SerLeuSerProTyrGlu 14

DB 65 TCTAACAACTGGCTGGCTTTCCTTTCACCGAACAACTCTTTCCTCTCTCATGAA 124

QY 15 GlnAsnHisHisArgLysAspValTyrSerSerThrThrThrValValAspValala 34

DB 125 TACAAC-----130

QY 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54

DB 130-----130

QY 55 PheProSerProPheGlyValValValAspAlaPheThrArgAspAsnSerHisSer 74

DB 131-----CTTGGCTTGGTCAGCGACCATATG-----GACAACTTTTCAAAACA 172

QY 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94
 DB 173 CAAGAGTGGATATG-----ATCAATCCACACGGTGGAGGAGGATGAA 217
 QY 95 -----ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThr 110
 DB 218 CGAGGAGAGGTTCAAAAGTGGCCGATTTCTCGTGTGAGCAAAACCGACGAAACCAA 277
 QY 111 AsnGluAsnValGlyAspGlySerGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGly 130
 DB 278 TCCAACCACTAGTAGCTTCAACAGCACTCAGACTACTTCTCCATATCACAAT-----328
 QY 131 GlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProVal-----147
 DB 329 -----AGCTTGATGCCTAGCGTCCAAACCAACGATCGTGTGTAGCAGCT 373
 QY 148 -----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
 DB 374 TGTGACTCCAATACTCTCAACAGTAGCTATCATGAGCTTCAAGAGAGTGTCTCACAAT 433
 QY 159 AlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
 DB 434 CTACAGTCACTACTTTTGTCCATGGGACCACC-----GCT 469
 QY 179 AsnAsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu---194
 DB 470 GGTATATATGTTGTAGACAAAGCTTCAACATCCGAGACCAACCGGGGATAACGTAGCGGT 529
 QY 195 -----AlaThrProLysLysThrIleGluSerPheGlyGln 206
 DB 530 GGAGCACTAGCCGTTGTTGAGACGGCCACGCAAGCGTGCATTCGACACTTCGGACAA 589
 QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226
 DB 590 CGAACTCGAUCTATCGTGTGTCAACAGACATCGATGCTGCTCGATATAGAGCTCAT 649
 QY 227 LeuTrpAspAsnSerCysLysArgGluGlnThrArgLysGlyArgGlnValTyrLeu 246
 DB 650 CTATGGGATAATAGTTGTAGAGGAGGAGCCAGTCTAGGAAAGGAGAGCAAGTTTACTTG 709
 QY 247 GlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyr 266
 DB 710 GGTGGATATGCAAAAGAGATAAAGCAGCAAGATCATATGATCTAGTGCATTAAGTAC 769
 QY 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
 DB 770 TGGGTCTCTCAACTACTACTAATTTCCCACTACGAGAAAGAGAGATAGAGGAA 829
 QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPhe 306
 DB 830 ATGAACACATGACGAGACAGAGATTCTGGTGGCTGCATTTAGAGGAAAGTATGTGATT 889
 QY 307 SerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGln 326
 DB 890 TCCAGAGCGCTTCGATGTATCGAGGAGTTACAAGGCATCACCACATCGAAGATGGCAA 949
 QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
 DB 950 GCAAGGATCGCGCGATCGCCGAAACAAAGACCTCTACTTCGGAACTTTTAGCTAGT 1009
 QY 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
 DB 1010 GAAGAAGCAGCAGAGCTTACGATATAGCTGCAATAAAGTTTAGAGGACTTAATGCAGTG 1069
 QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
 DB 1070 ACCAACTTCGAGATCAACCGGTACGCGTGAAGCCATTTCTAGAGTAGTACACTCTTCC 1129
 QY 387 Ile---GlySerAlaAlaLysArgLeuLysGluAla-----397
 DB 1130 ATCGAGGAGCGCGACCTAAACGGCTCAAGAGAGCTCAAGCTCTTGTAGTCTTCAAGAAA 1189

```

QY 397 ----- 397
Db 1190 CGCGAGCGGAGATAGATAGCCCTTGGTTCAAGTTTCAGTAGCGTGGTGGCTCGAGCACA 1249
QY 398 ----- 398
Db 1250 GGCTCTGGCTCCACCTCATCAGAGCTTCAGCTTCAACCTTACCCTCTAAGCATTCACAA 1309
QY 400 ProValProSerMetMetMetMetMetMetMetMetMetMetMetMetMetMetMetMetMet 419
Db 1310 CCATTAGAGCCCTTTCTCTCTCTCAGAACATGATCTCTCATTACACAAACAACAAT 1369
QY 420 GlyTyrGlnAsnAlaAlaValGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 438
Db 1370 GCTCAGATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1429
QY 439 GlnHis----- 445
Db 1430 CAACAGACCAACAATTACTTTCGAGCAACAGTCGAGCCAGAACTCTCAGCAGCTCTACAAT 1489
QY 446 GlyTyrTyrTyr----- 449
Db 1490 GGTATCTTCATAGCAATCCGGCTCTGCTTCATGAGCTTGTCTCTACCTCTATCGTTGAC 1549
QY 450 ----- 450
Db 1550 AACAATAATAACAATGGAGGCTCTAGTGGAGCTACACACTGCAGCA----- 1597
QY 466 GluAspAspGlnHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis 475
Db 1598 ----- 1598
QY 476 ----- 476
Db 1640 GTTGGATCGACCGAGGAGCTTCCCAACCGTTAAACAGATTACGATATGCTTCCAGTGAT 1699
QY 492 AspSerValThrValCysGlnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAla 511
Db 1700 ----- 1723
QY 512 ProValAsnCysAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArgAsnHisTyr 531
Db 1723 ----- 1723
QY 532 TyrPheAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 551
Db 1724 ----- 1753
QY 552 ThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGlyGlyGlyGlyGly 571
Db 1754 ----- 1759
QY 572 ThrPheThrValTrpAsnAsp 578
Db 1760 GTTTTCACTATGTGAATGAG 1780

RESULT 11
US-10-183-687-486
; Sequence 486, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Caboon, Rebecca
; APPLICANT: Ebelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime

```

```

; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BB1458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 486
; LENGTH: 1935
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; US-10-183-687-486

Alignment Scores:
Pred. No.: 2,58e-108 Length: 1935
Score: 1017.00 Matches: 266
Percent Similarity: 51.81% Conservatives: 63
Best Local Similarity: 41.89% Mismatches: 178
Query Match: 32.61% Indels: 128
DB: 13 Gaps: 24

US-09-980-364-2 (1-579) x US-10-183-687-486 (1-1935)

QY 3 AsnAsnTrpLeuGlyPheSerLeuSerPro-----TyrGluGlnAsnHisHisArgLys 20
Db 21 AACAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 74
QY 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
Db 75 AGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122
QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
Db 123 ACTTCGCGAGCC-----GCCGCCGTCCCAACAACAACACACACCCGCCA----- 164
QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTyrAspIleAsn 80
Db 165 -----AACTCTTATCAGCCGAAAAAGATGCAGGT 194
QY 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAsp-----Gly 94
Db 195 GCGGAAGGTCCAGCTACCAACCGCTAGTGCACCAACAGACTTGTCCATGTTTAAAGGTGCG 254
QY 95 ProLysLeuGluAsnPheLeuGlyArgThrThrThrThrThrThrThrThrThrThrThr 114
Db 255 CCGAACTCGAGATTTCTTGGCGGTCTCTCTACTACA----- 293
QY 115 GlyAspGlySerGlySerGlyCysTyrGlyGlyAspGlyGlyGlyGlyGlySerLeuGly 134
Db 294 -----CCTGGAGTTGGAGGTGCACCTCAACTGCAG 323
QY 135 LeuSerMetIleLysThrThrLeuArgAsnGlnProValAspAsnValAsnGlnGlu 154
Db 324 TTGCTTCATCATCACTTCTCTACTCAACCCCGTACTGTTTCTGCAACCATATATA 383
QY 155 AsnGlyAsnAlaAlaLysGlyLeuSer-----LeuSerMetAsnSerSerThrSer 171
Db 384 TATGATTCGGAGCTCAAAACCACTTCTCTATATTTCTTCTGGATTTCTCTCTGACCAA 443
QY 172 CysAspAsnAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThrIleAspAsp 191
Db 444 ATCGATATTCACAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 482
QY 192 SerValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyr 211
Db 483 -----GAGCCTGGACCCCAAAAGATCTGTGATCTTTGGCCCAACCTACTCTGATCTAC 536
QY 212 ArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTyrAspAsnSer 231
Db 537 CGGGGTGTCCACGAGCATAGATGCAAGGTATGAAGCTCACTTGTGGGCAACAATAGT 596

```

QY 232 CysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLys 251
Db 597 TGCAGAGAGAGAGGCGAAAGTAGGAAGGAGGCGAAGTTATTTCGGTGGCTATACACAG 656
QY 252 GluGluValAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyThrThr 271
Db 657 GAAGATTAAGCTGCAGAGCTTATGATCTTCGGCTCTCAAGTACTGGGTCGACCACT 716
QY 272 ThrThrAsnProMetSerGluTyrGluLysGluValGluGluMetLysHisMetThr 291
Db 717 ACTACCAACTTTCGGATTTCCAACTATGAGAGGAGCTCGAAGAGATGAAGAACATGATG 776
QY 292 ArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSer 311
Db 777 AGCAAGAGTTCGTTCTCTCCCGAGAGAAAGTAGTGATTTTCAGAGGTCCTTCA 836
QY 312 IleTyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgLysGlyArg 331
Db 837 ATTTCAGAGGAGTGAACGACATCATCAACATGGTAGGTGGCAAGCAAGAATTTGAAGA 896
QY 332 ValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyGlnGluAlaAlaGlu 351
Db 897 GTTGCAGGCAACAAAGATCTCTATCTTGGCACATTTAGCACCCCAAGAGAGAGCTGAA 956
QY 352 AlaTyrAspIleAlaAlaLysPheArgGlyLeuThrAlaValThrAsnPheAspMet 371
Db 957 GCCTATGATATTCAGCAATCAAGTTTAGAGGCTTAAACGGGTTTACAAATTTTGATAG 1016
QY 372 AsnArgTyrAsnValLysAlaLysLeuGluSerProSerLeuProIleGlySerAlaAla 391
Db 1017 AGCGGTACCATGATAACAGCATTTCAACAGC---AATCTTCCCATTTGGAGCACTGAAC 1073
QY 392 LysArgLeuLysGluAlaAsnArgProValProSerMetMetMetIleSerAsnVal 411
Db 1074 AACCAATCCAAACTGCAATGATTCACAT-----TCTGATAGGAGC 1115
QY 412 SerGluSerGluAsnSerAlaSerGly-----TrpGlnAsnAlaAla 425
Db 1116 ACCGATGAGGAGGAGATCATCTCTGGATCATCGGCGATGACCTTTACATCTCAGCCCGCT 1175
QY 426 ValGlnHisHisGlnGlyValAspLeuSerLeuHisGlnHisGlnGluArgTyrAsn 445
Db 1176 GCAAGCTCCAGTAGGCTTTGCAATTACATCAAGCAACCCATCCGATATTGGTCA 1235
QY 446 -----GlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCys 462
Db 1236 AACGTTCTCGATACATGATGATCTGCTGCTTCTTAAAGCAGT---GCCAAGAACCCCTATA 1292
QY 463 PheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMetThrAsn 482
Db 1293 TTTTCATCACTCACCAATGGATCGGTTTCCAAAGT---CCAAAGTGGATTCGTACCGGC 1349
QY 483 IleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnValValGly 502
Db 1350 AATCGTGGGATTCGAGTCAACGGA-----AGCATTAACATGGGTATTTCAT 1400
QY 503 TyrGlyGlyTyrGlnGly-----508
Db 1401 CGAGGTGGTATGTTGGGCGAGGAGCAACAGAGAGCAGTGTCTACTAATATCCCAACCCA 1460
QY 509 PheAlaAlaProVal-----AsnCysAspAlaTyrAlaAlaSerGluPheAspTyr 525
Db 1461 TTGGCAACGCCAATGCTTAAATATAGCGATTAATTAGAGCTGTCTCTCGGCTAT 1520
QY 526 AsnAlaArgAsnHisTyrTyrPheAlaGlnGln-----GlnGlnThrGlnGlnSerProGly 544
Db 1521 GGT-----GATACTGGAACTCGACAGTCGTTTCAATCTACTTATCAAACTGCAAG 1574
QY 545 GlyAspPheProAlaAlaMetThr-----552
Db 1575 CCAAGTCTCTCAGTGTTCCAACACCAACCAATTTTGGCATGGAATGAATCAGCTTCACATG 1634

QY 553 ---AsnAsnValGlySerAsnMetTyrTyrHisGly-----GluGly 565
Db 1635 GACATTAACCTGTGAAAAAATCTG---TGGAGGCGCGATGCATTCAGAGTTGGAAGG 1691
QY 566 Gly-----GlyGluValAlaProThr 572
Db 1692 GGTAGGTTGTTTAACTTTCTTGAATAGGTGAGATTGGAATGACA 1736
RESULT 12
US-10-183-687-484
; Sequence 484, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 484
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Catalpa sp.
US-10-183-687-484
Alignment Scores:
Pred. No.: 1,15e-106 Length: 1959
Score: 1003.00 Matches: 238
Percent Similarity: 54.20% Conservative: 46
Best Local Similarity: 45.42% Mismatches: 84
Query Match: 32.16% Indels: 156
DB: 17 Gaps: 17
US-09-980-364-2 (1-579) x US-10-183-687-484 (1-1959)
QY 2 AsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLysAsp 21
Db 109 CATCAGAACTGCTCGCTTCTCGCTTTCC-----AACCATCT-----147
QY 22 ValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAspPro 41
Db 148 -----CTGTTTCAACCC 159
QY 42 ThrAlaAlaSerAspGluSerSerSerAlaIleGlnThrSerPheProSerProPheGlyVal 61
Db 160 TCGTTCAACTCTAATCCCTCCCTCCAGCTGC-----GGGTT 195
QY 62 ValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsnGly 81
Db 196 GTG-----AATCAAGAAATGCGGT-----GCGGT 222
QY 82 CysAlaCysAsnAsnIleHisAsnAsp-----GluGlnAspGlyProLysLeuGluAsn 99
Db 223 GGTCCGCAACAGATCATCTCTCGGTTCTTTCAGTGGAGGCGCAAGCTGGAGAT 282
QY 100 PheLeuGlyArgThrThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
Db -----

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ. ID NOS: 285684
 ; SEQ. ID NO 100930
 ; LENGTH: 2116
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_62152C.1
 US-10-424-599-100930

Alignment Scores:

Pred. No.: 8,54e-106 Length: 2116
 Score: 996.00 Matches: 233
 Percent Similarity: 61.59% Conservative: 62
 Best Local Similarity: 48.64% Mismatches: 107
 Query Match: 31.93% Indels: 77
 DB: 13 Gaps: 14

US-09-980-364-2 (1-579) x US-10-424-599-100930 (1-2116)

QY 2 AsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLysAsp 21
 DB 275 TCGACTAACTGGGTATCGTTCTCTCTCCCATGGAATGCTCGAACCCTCGAACCT 334
 QY 22 ValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAspPro 41
 DB 335 CAGTTCCTTCAA-----TACGAC--- 352
 QY 42 ThrAlaAspSerAspGluSerSer-----AlaIleGlnThrSerPheProSerPro 58
 DB 353 ---GCCGCTTCGACACTTCTCGCATCACTACTCTCGACAACTGTGACCAACGGG 409
 QY 59 PheGlyValValValAlaAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAsp 78
 DB 410 TGGGG-----AACGGGAGCCTCAAGTTTGAGCAGAAAT 442
 QY 79 IleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAsp-----Gly 94
 DB 443 CTGAACACACGACGACGTGATTCGTGTGAATCGTCGCGAGAGCGTCAGCCACCGCG 502
 QY 95 ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsn--- 113
 DB 503 CCGAAGCTGGAGGATTTCTCGGCGACTCTCGGCTGTTATGCGTTACTCCGACACCGAG 562
 QY 114 -----ValGlyAspGlySer----- 118
 DB 563 ACGGAGACGACGACTCGTCGCTGACGCACATCTAGACACACACACACACACAC 622
 QY 119 -----GlySerGlyCysTyrGlyGlyAspGlyGly-----GlySer 132
 DB 623 CACCACACACACGCTTCTTCGCTACTTCGGGGTGACACACACACACATCTCAAGGCCATT 682
 QY 133 LeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsnValAsp-As 152
 DB 683 ACTGGATTCGAAGCTTTTCGACTTAACCTCTGGCTCCGAGGTTGATTTCTGCAATC 742
 QY 152 nGlnGluAsnGlyAsnAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCy 172
 DB 743 GGAAGAGCGGACGACGAGTTCGGGACTCACTCTATTGAGTCTCCGCGTGGCCTCGGATTC 802
 QY 172 sAspAsnAsnAspSer-AsnAsnAsnValValAlaGlnGlyLysThr---Ile-Asp 190
 DB 803 GCCGCGTTCTCCGGTGGCAGACACCGAACCTTGTGCTCGCGTGGCACTGAGCTCCGAG 862
 QY 191 AspSerValGluAla-----ThrProLysLysThrIleGluSerPheGlyGln 206
 DB 863 AAGCCCGTTGCGCGCGGAGTCCAAATAGTTCGAAGAAGATCGTGGATACCTTCGGCCAG 922
 QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226

DB 923 CGGACTTCTATTACAGAGGTGTTACTAGGCACCGATGACAGGAAGATATGAAGCGCAT 982
 QY 227 LeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeu 246
 DB 983 CTAATGGGACATAGTTTGCAGAGAGGAGGTCAAGCCAGAAAGGCGCTCAAGTTATTG 1042
 QY 247 GlyGlyTyrAspLysGluGlyLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyr 266
 DB 1043 GGTGGATATGATAAGGAAGAAAGGCCCGCAGAGCTTATGATTTGGCAGCTCTAAGTAC 1102
 QY 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
 DB 1103 TGGGTCCTCCACTGTACCCCACTTCCCTGTTTCCCAATTATTCGAGGAGTGGAGAG 1162
 QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgLysSerSerGlyPhe 306
 DB 1163 ATGAACATGTAAACAAGCAAGAATTATCGCATCATTCGCGAGGAAAAGTAGTGGTTTC 1222
 QY 307 SerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGlyArgTrpGln 326
 DB 1223 TCCAGGGAGCTTCCATATACAGAGGTGTAAAGGCATCATCAACAGGTAGGTGGCAA 1282
 QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
 DB 1283 GCAAGAAATGGCCGTGTAGCTGGAACAAGATTTATACTTGGGAACATTGCAACCGAG 1342
 QY 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaLysPheArgGlyLeuThrAlaVal 366
 DB 1343 GAGGAAGCAGCAGAGGCATATGATATTCAGCCATAAAGTTTCAGAGGTGCAACCGGTA 1402
 QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
 DB 1403 ACCAACTTTGAGATGATAGATATGATGTGGAGCTATATGAAGAGT---TCTCTTCCA 1459
 QY 387 IleGlySerAlaAlaLysArgLeuLys-----GluAlaAsnArgProValProSer 403
 DB 1460 GTGGGTGGGCGACGCAAAACGCTTGAGGCTTTCCCTTGAATCAGACGACGAAAGCTCTCT 1519
 QY 404 MetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGlyTrpGlnAsn 423
 DB 1520 GTG-----AACAGCAGCAGTCAGCAGAGAT 1546
 QY 424 AlaAlaValGlnHisHisGlnGly---ValAspLeuSerLeuHisGln 439
 DB 1547 CCACAGTGTGGTAAAGTGGTAGCATCAATTTCTCAGCCATTCATCAG 1597

Search completed: July 19, 2004, 20:17:48

Job time : 809 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 09:59:14 ; Search time 20 Seconds
(without alignments)
2784.745 Million cell updates/sec

Title: US-09-980-364-2
Perfect score: 3119
Sequence: 1 MNWLGFLSPYEQNHHRK.....YVHGEGGGEVAFVFWNDN 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2475	79.4	581	2	T51580	ovule development
2	1040.5	33.4	516	2	F96549	hypothetical prote
3	997	32.0	566	2	T49988	ovule development
4	930	29.8	485	2	T03638	hypothetical prote
5	926	29.7	425	2	B96750	hypothetical prote
6	891	28.6	555	2	S71365	ovule development
7	654.5	21.0	332	2	D86295	hypothetical prote
8	619.5	19.9	308	2	H96827	protein F20B17.12
9	572	18.3	485	2	S54116	hypothetical prote
10	511	16.4	205	2	T47591	aintegumenta-like
11	475.5	15.2	433	2	T01574	indeterminate spik
12	465.5	14.9	464	2	C84686	probable AP2 domai
13	461.5	14.8	432	2	B85436	APETALA2 protein
14	459	14.7	436	2	B84845	probable AP2 domai
15	446	14.3	446	2	T03981	APETALA2-like prot
16	270	8.7	247	2	T06721	hypothetical prote
17	269	8.6	236	2	T02577	probable AP2 domai
18	197	6.3	292	2	T09030	hypothetical prote
19	180	5.8	358	2	D96579	hypothetical prote
20	162	5.2	314	2	E86452	protein FSU5.5 [m
21	157.5	5.0	277	2	T03927	DNA binding protei
22	155.5	5.0	336	2	B84594	AP2 domain transcr
23	153.5	4.9	207	2	T49897	transcription fact
24	150	4.8	171	2	T00432	probable AP2 domai
25	150	4.8	245	2	B86168	hypothetical prote
26	149.5	4.8	328	2	G86263	hypothetical prote
27	149	4.8	204	2	B86410	protein F3M18.20 [
28	148.5	4.8	261	2	B86354	hypothetical prote
29	148	4.7	245	2	B86407	hypothetical prote

ALIGNMENTS

RESULT 1

T51580

ovule development protein aintegumenta-like protein - Arabidopsis thaliana
N:Alternate names: protein T10B6_90

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000

C:Accession: T51580

R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew

submitted to the Protein Sequence Database, August 2000

A:Reference number: Z25394

A:Accession: T51580

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-581 <SAT>

A:Cross-references: EMBL:AL391142

A:Experimental source: cultivar Columbia; BAC clone T10B6

C:Genetics:

A:Map position: 5

A:Introns: 78/1; 216/2; 244/1; 273/3; 298/2; 315/2; 341/1

A:Note: T10B6_90

Query Match 79.4%; Score 2475; DB 2; Length 581;

Best Local Similarity 81.3%; Pred. No. 2.3e-166;

Matches 474; Conservative 36; Mismatches 59; Indels 14; Gaps 7;

QY	3	NNWLGFLSPYEQNHHRKDVYSSTTTTVVDVAGEYCYDPTAASDESSAIOISFPSPFGV	62
DB	5	NNWLGFLSPHDQNHRTDVDSSTRTAVDVAGGYCFDLAAPSDESSAVQTSFLSPFGV	64
QY	63	VDATRDNNSHSRDWDINGACNNIHNDQDQPKLENFLGRTTTTYNTNENVDGSGSGC	122
DB	65	LEAFTRDNNSHSRDWDINGACNNINNEQNGPKLENFLGRTTTTYNTNETVVDGNGD-C	123
QY	123	YGGGPGGGGSLGLSMIKTWLRNCPVDVNDQNGNAAGLSLSMNSSTSCDNNNSNNV	182
DB	124	GGDGGGGGSLGLSMIKTWLSNVANHQNGARGLSLSMNSSTSCDNNNNDD	182
QY	183	VAQGTIDDSVEATPKKTIESFGQTSYRGVTRHRTGRVEAHLWDNSCKREGQTRKR	242
DB	183	VVOETIVDVWETTPKKTIESFGQTSYRGVTRHRTGRVEAHLWDNSCKREGQTRKR	242
QY	243	QYVLGGYDKEEKAAYDLAALKYWGTTTTFPMSEYEKEVEEMKHTROYVASLRK	302
DB	243	Q---GGYDKEEKAAYDLAALKYWGTTTTFPMSEYEKEVEEMKHTROYVASLRK	299
QY	303	SSGFGSGASIRGVTRRHQHQWQARIQRVAGNKDLYLGTGTQBEAAEAYDIAAKFRG	362
DB	300	SSGFGSGASIRGVTRRHQHQWQARIQRVAGNKDLYLGTGTQBEAAEAYDIAAKFRG	359
QY	363	LFAVTFDNNRVNVKALLESPLPGCSAAKRLKEANRPVPSMMISNNVSSSENSASGW	422
DB	360	LSAVTFDNNRVNVKALLESPLPGCSAAKRLKEANRPVPSMMISNNVSSSENSASGW	418

QY 423 NAAVQHGHQVDSLIIHQHQRERYNGYVYNGNLSSESARACFKQEDDQHHFLSNTQSLMTN 482
 Db 419 NTAQHHQGMDSLIIQQQERYVG-YVYNGNLSSTESTRVCFKQEEQHHFLRNSPESHTN 477
 QY 483 IDHSSVSDSDSVTCGVVGGVGGQGAAP-----VNCDAVAASEFDYNARHHYFA---Q 535
 Db 478 VDHHSSTSDSDSVTCGVVGGVGGQGAAP-----VNCDAVAASEFDYNARHHYFA---Q 537
 QY 536 QOQOQSGGDFPAAMTNNVGSNNYHGGEGCEVAPFTVWMD 578
 Db 538 QOQOQSGGDFPAAMTNNVGSNNYHGGEGCEVAPFTVWMD 580

RESULT 2
 F96549
 N;Alternate names: protein F11M15.6 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: F96549
 R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; PMID:21016719; PMID:11130712
 A;Accession: F96549
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-516 <STO>
 A;Cross-references: GB:AB005173; NID:94836931; PIDN:AAD30633.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: F11M15.6
 A;Map position: 1

Query Match 33.4%; Score 1040.5; DB 2; Length 516;
 Best Local Similarity 44.6%; Pred. No. 1.7e-65;
 Matches 254; Conservative 70; Mismatches 114; Indels 131; Gaps 22;

QY 86 NIHNDQDG---PKLENFLGRITTT-IYNTNEN-----VGDGSGGCGYGGDGGGSLGL 135
 Db 3 NPHGGGEGGEPKVPADFLGVSKGDHHTDNLVYPYNDIHQTNASDYFQTN---SL-L 57

QY 136 SWIKTWLRNQPVDVNDQENGNAAKGLSLSNSS-----TSCDNNDSN 179
 Db 58 PTVTTCASNAP--NNYLOESAHNLQSLTSLMSGTCAAAAEVATVKASPAETSDNSSTT 116

QY 180 NNVAQGTIDDSVEATPKTIESFGORTSYRGVTRHRTGTYEAHLWNSCKREGQTR 239
 Db 117 N---TSGGAI---VEATPRRTLETFGQRTSYRGVTRHRTGTYEAHLWNSCKREGQSR 170

QY 240 KGRQVYLGYYDKEEKAARAYDLAALKYWGTTTNTTTPMSYEKEVEEMKMTQEVYASL 299
 Db 171 KGRQ---GGYDKEEKAARAYDLAALKYWGTPSTTTNFTITNYEKEVEEMKMTQEVASI 227

QY 300 RKSSGFGSGASTYRGVTRHRTGTYEAHLWNSCKREGQTR 359
 Db 228 RKSSGFGSGASTYRGVTRHRTGTYEAHLWNSCKREGQTR 287

QY 360 FRGLTAVTNFPMNRYNKALLESPLI-GSAAKRLKEA-----NRPVPMNRYNKALLESPLI 413
 Db 288 FRGLTAVTNFPMNRYNKALLESPLI-GSAAKRLKEA-----NRPVPMNRYNKALLESPLI 347

QY 414 SENSASGQNA-----VOHQGV-----DLS----- 435
 Db 348 -YGAASGSSVASSRRLQYPYLSIQOQFPHLHHQPLTLQNNNDISQVHDSFSVIQT 406
 QY 436 LLAHQHQRERYNGYVYNGNLSSESARACFKQEDDQHHFLSNTQSLMTNIDHQSVDSDSVT 495

Db 407 QULHHQOQTNNYLOSSHTSQL-----YNAYLOSNPGLL-----HGFVSDNN-- 448
 QY 496 VCGNVVGYGGYQGAAPVNCDAVAASEFDYNARHHYFAQQOQOQTSQSGGDFPAAATNNV 555
 Db 449 ---NTSGFLGNGGIGSSSTVGSAAEFPPAVKVDY-----DWPSPS----- 487
 QY 556 GSNMYHGGEGCEVAP-----TFTVWMD 578
 Db 488 GGATGYGWNSSGESAQSGNPGGVFTWNE 516

RESULT 3
 T49988
 N;Alternate names: protein-like - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C;Accession: T49988
 R;Bevan, M.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, April 2000
 A;Reference number: Z25026
 A;Accession: T49988
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-566 <BEV>
 A;Cross-references: EMBL:AL353995; GSPDB:GN00063; ATSP:F12B17.140
 A;Experimental source: cultivar Columbia; BAC clone F12B17
 C;Genetics:
 A;Gene: ATSP:F12B17.140
 A;Map position: 5
 A;Introns: 33/1; 259/2; 287/1; 316/3; 341/2; 384/1

Query Match 32.0%; Score 997; DB 2; Length 566;
 Best Local Similarity 40.4%; Pred. No. 2.2e-62;
 Matches 247; Conservative 58; Mismatches 127; Indels 180; Gaps 18;

QY 43 AASDESSAIQTSFSPFPFGVWVDAFTRDNNSHSRDWDINGCACNNIHNDQDQPKLENFLG 102
 Db 52 AASMDSTILITTFVDP-----QSHHSQNHIPKLEDFLG 84

QY 103 RT-----TTIY-----NTNEN 113
 Db 85 DSSSIVRSDNSQTDQDSSLTQIYDPRHHNQTSFYSDHDFDKTMAGFQSAFSTNSGSE 144

QY 114 VGDGSGSG-----CYGGDGGGSGLSMIKTLWRNQPVDVNDQ 153
 Db 145 VDSASIGRTHLAGDYLGHVVESSGPELGFHGGSTGALSGLGVNWN--TNHRDN-DNH 201

QY 154 ENGNAAKGLSLSNSSSTSCDNNDSNNNVVAGK-----TIDDSVEATPKTIESFGORTS 209
 Db 202 YRGN-----NNGERINNNNDNEKTDSEKFAVVAVVETSDCSNKKIADTFGORTS 252

QY 210 IYRGVTRHRTGTYEAHLWNSCKREGQTRGROYVYLGYYDKEEKAARAYDLAALKYWG 269
 Db 253 IYRGVTRHRTGTYEAHLWNSCKREGQTRGROYVYLGYYDKEEKAARAYDLAALKYWG 309

QY 270 TTTTNPMSYEKEVEEMKMTQEVYASLRRKSSGFGSGASTYRGVTRHRTGTYEAHLWNSCK 329
 Db 310 TATTNFTITNYEKEVEEMKMTQEVYASLRRKSSGFGSGASTYRGVTRHRTGTYEAHLWNSCK 369

QY 330 GAVAGNKDLYLGTGTQTEFAAEAYDIAIKRGLTAVTNFDMNRYNKALLESPLIPTS 389
 Db 370 GAVAGNKDLYLGTGTQTEFAAEAYDIAIKRGLTAVTNFDMNRYNKALLESPLIPTS 428

QY 390 AAKRLKEANRPVPMNRYNKALLESPLI-GSAAKRLKEA-----NRPVPMNRYNKALLESPLI 449
 Db 429 AAKRLK-----LSLEAAASSEQKFIHLGHQ-----LHFFQOQO----- 461

QY 450 NCGNLSSESARACFKQEDDQHHFLSNTQSLMTNIDHQSVDSDSVTCGVVGGYGGV 509
 Db 462 -----QOQOQLOQSSPNHSSINFALCPNSAVOSQOQIIPCG----- 496

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: AB6141; MUID:21016719; PMID:11130712
A;Accession: D86295
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <STO>
A;Cross-references: GB:AE005172; NID:g6587812; PIDN:AAF18503.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 21.0%; Score 654.5; DB 2; Length 332;
Best Local Similarity 51.5%; Pred. No. 1.3e-38;
Matches 135; Conservative 42; Mismatches 66; Indels 19; Gaps

QY 167 NSTSCDN---NNDSNNV--AQKTTDDSVETPKKTIESFGORTSIYGVTRHWGTG 222
 :
Db 8 SKKTSVENETGGDQATSVLKKRKRQRDPAPP-----QRSVHGRVTIRHWGTG 59
 :
QY 222 RVEAHLWNSCKREGQTRKGROYLVGGYDKKEAKARAYDLAALKYWGTTTTTFPMSEYE 288
 :
Db 60 RVEAHLWDKNSWNETQTCKGRQ--GAYDEEDAARAAYDLAALKYWGSDTILNFLCNYE 111
 :
QY 282 KVEEMKHMTROEYVASLRKSSGFSGRCASYRGVTRHHQHGRQVARGVAGNKDIYL 348
 :
Db 117 EDIKEMESQSKEEYIGSLURKSSGFSGRVGYGVAKHHNGWEARIGRVFGNKIYL 178
 :
QY 342 TFGTOEEAAEYDIAAIKFRGLTAVTNFDMRYNVKAILESPSIFIGSAARKLEARNPV 403
 :
Db 177 TVATOEEAAIAYDIAAIEYGLNAVTFNFDISKY--LKLVPENPIDTANNLLSPHSDL 233
 :
QY 402 PSMIMISNNVSESENSASGWQN 423
 :
Db 234 SPFIKPHESDLQSQSSSEDN 255
 :

RESULT 8
H96827
Protein P20B17.12 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: H96827
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, C.
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maici, R.
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: AB6141; MUID:21016719; PMID:11130712
A;Accession: H96827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-308 <STO>
A;Cross-references: GB:AE005173; NID:g7715603; PIDN:AAF68121.1; GSPDB:GN00141
C;Genetics:
A;Gene: F20B17.12
A;Map position: 1

Query Match 19.9%; Score 619.5; DB 2; Length 308;
Best Local Similarity 58.4%; Pred. No. 3.5e-36;
Matches 125; Conservative 31; Mismatches 41; Indels 17; Gaps

QY 206 QRTSYRGVTRHWTCGRYEALWNDSCKREGQTRKGROYLVGGYDKKEAKARAYDLAALK 268
 :
Db 48 QRSSFYRGVT--RWITGRYEALWDKNNSWNTDTTKGRQ--GAYDEEAARAAYDLAALK 103
 :
QY 266 YGWTITTTFPMSEYEKEVEEMKHMTROEYVASLRKSSGFSGRCASYRGVTRHHQHGRW 328
 :
Db 103 YNGRDTLINFLPYSYDEDKMEQGSKKEEYIGSLURKSSGFSGRVGYGVARHHNGRW 161

T47591

0v 224 EAH.WDNSCKREGOTRKGROVY.GGYDKEEKAARAYDI.AA.T.KVWGTTTTTTFPMSEYEKE 283

DD 00 FAEFEFEFVFAHFVFVVMQFQNHEDDUGNTAQNFVAFANVIRKOFKNSNSQYIRKSVIFIANIOWM 123

DD 00 FAEFEFEFVFAHFVFVVMQFQNHEDDUGNTAQNFVAFANVIRKOFKNSNSQYIRKSVIFIANIOWM 123

Search completed: July 15, 2004, 10:03:41
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 09:52:29 ; Search time 13 Seconds
(without alignments)
2319.125 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119
Sequence: 1 MNNWLGFLSPYEQNHRRK.....YHGGCGGEVAFPTFWNDN 579

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	451.5	14.8	432	1 AP2_ARATH	P47927 arabidopsis
2	150	4.8	248	1 AP23_ARATH	P4736 arabidopsis
3	145	4.6	248	1 PT16_LYCES	O04682 lycopersico
4	135.5	4.3	1146	1 AS10_YEAST	P48361 saccharomyc
5	132.5	4.2	344	1 RAV1_ARATH	Q92wm9 arabidopsis
6	132	4.2	243	1 ERF2_ARATH	O80338 arabidopsis
7	129.5	4.2	161	1 PT15_LYCES	O04681 lycopersico
8	126.5	4.1	1291	1 VAC4_HELPY	Q48258 hellobacte
9	125	4.0	222	1 ERF4_ARATH	O80340 arabidopsis
10	124	4.0	268	1 ERF1_ARATH	O80337 arabidopsis
11	123	3.9	1341	1 YL78_YEAST	Q05854 saccharomyc
12	122.5	3.9	1253	1 DSPP_HUMAN	Q9nzv4 homo sapien
13	122	3.9	2339	1 RPL1_PLAFA	P27625 plasmodium
14	121.5	3.9	1296	1 VAC1_HELPY	Q48247 hellobacte
15	121	3.9	961	1 GPR1_YEAST	Q2361 saccharomyc
16	120	3.8	964	1 CPA2_BACCI	P70873 bacillus ci
17	119.5	3.8	623	1 PNT1_DROME	P51022 drosophila
18	116.5	3.7	918	1 YMJ3_CABEL	P34487 caenorhabdi
19	116	3.7	1093	1 SM14_YEAST	P25302 saccharomyc
20	115.5	3.7	880	1 BCP_CABEL	Q24206 saccharomyc
21	114.5	3.7	1317	1 GRP_CABEL	P34288 caenorhabdi
22	114.5	3.6	1603	1 VIT3_CABEL	Q9n432 caenorhabdi
23	113.5	3.6	1603	1 VIT4_CABEL	P18947 caenorhabdi
24	113	3.6	300	1 ERF5_ARATH	O80341 arabidopsis
25	112.5	3.6	1290	1 VACA_HELPY	P55981 hellobacte
26	112	3.6	719	1 NRPI_YEAST	P32770 saccharomyc
27	112	3.6	865	1 HTR3_ECOLI	P33129 escherichia
28	112	3.6	1310	1 VAC3_HELPY	Q48253 hellobacte
29	111.5	3.6	1603	1 VIT5_CABEL	P66125 caenorhabdi
30	111.5	3.6	1912	1 VITI_CHICK	P87498 gallus gall
31	111	3.6	813	1 TTKA_DROME	P42282 drosophila
32	110.5	3.5	225	1 ERF3_ARATH	O80339 arabidopsis
33	110.5	3.5	416	1 FILA_HUMAN	P20930 homo sapien

34	110	3.5	405	1 YCC8_YEAST	P25367 saccharomyc
35	110	3.5	578	1 PSF2_YEAST	P50109 saccharomyc
36	110	3.5	955	1 YCGV_ECOLI	P76017 escherichia
37	109	3.5	753	1 CKAA_BACUF	O32321 bacillus th
38	109	3.5	943	1 LBPA_NEIMB	Q06379 neisseria m
39	109	3.5	1087	1 POM1_SCHPO	Q09690 schizosacch
40	109	3.5	1541	1 IGA1_HAEIN	P42782 haemophilus
41	108.5	3.5	730	1 GLN3_YEAST	P18494 saccharomyc
42	108.5	3.5	1113	1 N116_YEAST	Q02630 saccharomyc
43	108.5	3.5	1211	1 BUN2_DROME	Q24523 drosophila
44	108	3.5	164	1 P202_LUPPO	P16146 lupinus pol
45	107	3.4	547	1 YCGV_ECO57	Q8X583 escherichia

ALIGNMENTS

RESULT 1

AP2_ARATH	STANDARD;	PRT;	432 AA.
AC P47327			
DT 01-FEB-1996 (Rel. 33, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DT 10-OCT-2003 (Rel. 42, Last annotation update)			
DE Floral homeotic protein APETALA2.			
GN AP2 OR AT4G36920 OR AP22.49 OR C7A10.440.			
OS Arabidopsis thaliana (Mouse-ear cress).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.			
OX NCBI_TaxID=3702;			
RN [1]			
RP SEQUENCE FROM N.A., AND MUTANTS AP2-1 AND AP2-5.			
RC STRAIN=cv. Columbia; TISSUE=Flower;			
RX MEDLINE=95003709; PubMed=7919989;			
RA Jofuku K.D., den Boer B.G.W., van Montagu M., Okamoto J.K.;			
RT "Control of Arabidopsis flower and seed development by the homeotic			
RT Gene APETALA2.";			
RL Plant Cell 6:1211-1225(1994).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=cv. Columbia;			
RX MEDLINE=20083488; PubMed=10617198;			
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,			
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,			
RA Harris B., Ansoerge W., Brandt P., Grivell L.A., Rieger M.,			
RA Weichselgartner M., de Simone V., Obermaler B., Mache R., Mueller M.,			
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,			
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,			
RA Vos P., Hohnel J., Zimmermann W., Wedler H., Ridley P.,			
RA Langham S.-A., McCullagh B., Bilham L., Kobben J.,			
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,			
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,			
RA Weitzenecker T., Bothe G., Ramspeiger U., Hilbert H., Braun M.,			
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,			
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,			
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,			
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,			
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,			
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,			
RA Pettitt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,			
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,			
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,			
RA Gabel C., Fuchs M., Fartmann B., Grandterth K., Dauner D., Herzl A.,			
RA Neumann S., Argiriou A., Vitale D., Liguori K., Piravandi E.,			
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,			
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,			
RA Chedor F., Cooke R., Berger C., Monfort A., Casacuberta E.,			
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,			
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,			
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Balke C.,			
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker M.,			
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,			

FT	DNA BIND	220	286	AP2/ERF 2.	
FT	MUTAGEN	159	159	G->E: TEMPERATURE SENSITIVE; AP2-5.	
FT	MUTAGEN	251	251	G->S: TEMPERATURE SENSITIVE; AP2-1.	
FT	MUTAGEN	420	420	Q->E: TEMPERATURE SENSITIVE; AP2-5.	
SEQ	SEQUENCE	432 AA;	47833 MW;	376A355291EFPB42 CRC64;	

Query Match	14.8%;	Score	461.5;	DB 1;	Length	432;
Best local Similarity	34.3%;	Pred.	NO. 5.8e-24;			
Matches 114;	Conservative	59;	Mismatches	112;	Indels	47;
Gaps	8;					

QY	198	KTITSEFGCGT	SYRGVTRHRTGR	VEAHLWNS	CKREG	TRKRG	OVYLG	GYDK	EEKAAR	257
		:	:	:	:	:	:	:	:	:
DB	119	KXSRGPRSR	SGYRGVTR	YRTGR	WESHND	-C	-----	GK	YILGG	DTAHAAR
		:	:	:	:	:	:	:	:	:
QY	258	AYDLAKY	CGTTTTNF	PMSEYE	KEVE	EMKMT	QYV	VASLR	KSSG	PSRGS
		:	:	:	:	:	:	:	:	:
DB	169	AYDRAA	KFRG	VEAD	INF	IND	YDD	DD	LKQ	MTLKEE
		:	:	:	:	:	:	:	:	:
QY	318	RHHCHGR	QWAR	IGR	VAGN	KDLV	LG	FTG	CEAA	EYDIAA
		:	:	:	:	:	:	:	:	:
DB	229	L-HKCGR	WAR	QGL	FLG	KYV	ILG	FD	TSVE	ARA
		:	:	:	:	:	:	:	:	:
QY	378	AILES	PLPG	SAAK	ELKE	ANRP	VS	MMMI	SNNV	SE
		:	:	:	:	:	:	:	:	:
DB	286	LNAES	-----	SG	NPTT	PQ	HN	LD	LS	GN
		:	:	:	:	:	:	:	:	:
QY	431	-----	GVD	LSL	HQ	HER	YNG	YV	NGN	LS
		:	:	:	:	:	:	:	:	:
DB	335	SNEV	LG	LT	GT	MLNT	TP	NS	HN	QF
		:	:	:	:	:	:	:	:	:
QY	484	-----	DH	QSV	SD	SV	TC	NV	CG	NV
		:	:	:	:	:	:	:	:	:
DB	394	AA	SSG	FSP	HH	QIF	N	ST	ST	P
		:	:	:	:	:	:	:	:	:

RESULT 2	AP23	ARATH			
ID	AP23	ARATH	STANDARD;	PRT;	248 AA.
AC	P42736;	Q23105;			
DT	01-NOV-1995	(Rel. 32,	Created)		
DT	28-FEB-2003	(Rel. 41,	Last sequence update)		
DT	15-MAR-2004	(Rel. 43,	Last annotation update)		
DE	AP2 domain transcription factor RAP2.3	(Related to AP2 protein 3)			
DE	(Cadmium-induced protein AG30).				
DE	AT3G16770 OR MGL6.1 OR MGL6.24.				
OS	Arabidopsis thaliana (Mouse-ear cross).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;				
OC	eurosid1 I; Brassicales; Brassicaceae; Arabidopsids.				
NCBI_TaxID=3702;					
[1]	[1]				
RN	SEQUENCE FROM N.A.				
RN	STRAIN=cv. Columbia;				
RC	Choi S., Baek E., Lee S.;				
RL	Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.				
[2]	[2]				
RN	SEQUENCE FROM N.A.				
RX	MEDLINE=97338147; PubMed=9192694;				
RA	Okumuro J.K., Caster B., Villarreal R., van Montagu M., Jofuku K.D.;				
RT	"The AP2 domain of APTAL2 defines a large new family of DNA binding				
RT	proteins in Arabidopsids.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 94:7076-7081(1997).				
[3]	[3]				
RN	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Columbia;				
RX	MEDLINE=20277480; PubMed=10819329;				
RA	Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;				
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence				
RT	features of the regions of 4,504,864 bp covered by sixty P1 and TAC				
RT	clones.";				
RL	DNA Res. 7:131-135(2000).				
[4]	[4]				
RN	SEQUENCE FROM N.A.				
RP					

RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Akrawa T., Bann J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallenberg E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinoraki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome."; [5]
 RL Science 302:842-846 (2003).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Brover V., Trukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.A.;
 RT "Full-length cDNA from Arabidopsis thaliana";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- INDUCTION: By cadmium.
 CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
 CC -!- CAUTION: Ref.1 sequence is most probably a cloning artifact and
 CC differs from that shown due to an insertion into the sequence.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z37504; CAA85734.1; ALT SEQ.
 DR EMBL; AF003096; AAC49769.1; -
 DR EMBL; AB022117; BAB02769.1; -
 DR EMBL; AY059917; AAL24399.1; -
 DR EMBL; AY035100; AAK59605.1; -
 DR EMBL; AY142562; AAN13131.1; -
 DR EMBL; AY087488; AAN65031.1; -
 DR HSSP; O80337; 2GCC.
 DR TRANSFAC; T02658; -
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PR00367; ETHRSPELEMT.
 DR PRODOM; PD001423; TF_ERF; 1.
 DR SMART; SM00390; AP2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein; Cadmium.
 FT DNA BIND 76 141
 FT DOMAIN 140 143 POLY-PRO.
 FT SEQUENCE 248 AA; 27758 MW; 317D50CC514782F1 CRC64;
 SQ
 Query Match 4.8%; Score 150; DB 1; Length 248;
 Best Local Similarity 26.1%; Pred. No. 0.0027;
 Matches 68; Conservative 38; Mismatches 105; Indels 50; Gaps 10;
 YQ 240 KGRVYLGYYDKERKAAAYDLAALKYWGTTTTFN-PMSEYEKEVEEMKHMTREQEYVAS 298
 DB 18 KGRKLTAEELWSELDAASADD-----PWGYSYSKLHPTNQVNVKBAVKEQATPEGKR 72
 YQ 299 LRRKSSGFSRGASTYRGVTRHQHGRWCARIGRVAGNKDLYLGTGFGQEEAAAYDAIAI 358
 DB 73 RKRK-----NVYRGI-RKRPWKAAEIRDPRKGVRLVGLGTFTAAEAAAYDVAAK 123
 YQ 359 KFRGIGTAVTFN-DMNRVNVKALLESPLPGSAAKRLKEANRPVPSMMISNNVSESENS 417
 DB 124 QIRGDKAKLNPPDLJHPPPPNYTPPPSP-----RSTDQP-PAKKVCVQSQSELS 174
 YQ 418 ASGWQNAVQHQQGVDSLHLLHQHCRYNGYNGYNGNLSESARACFRQEDDQHFLSNTQ 477

DB 175 QPSFVPECIGFGNGDEFQ-----NLSYG-----FEPDYDLKQKQISSLE 212
 QY 478 SLMT---NIDHQSSVSDSVT 495
 DB 213 SFLELDGNTAEQPSQLDESVS 233
 RESULT 3
 PT16_LYCSES
 ID PT16_LYCSES STANDARD; PRT; 248 AA.
 AC O04682;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE pathogenesis-related genes transcriptional activator PT16.
 GN PT16.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97357308; PubMed=9214637;
 RA Zhou J., Tang X., Martin G.B.;
 RT "The Pto kinase conferring resistance to tomato bacterial speck
 RT disease interacts with proteins that bind a cis-element of
 RT pathogenesis-related genes";
 RL EMBL J. 16:3207-3218 (1997).
 CC -!- FUNCTION: Transcription factor that binds to the GCC-box
 CC pathogenesis-related promoter element. Activates plants defense
 CC genes.
 CC -!- SUBUNIT: Interacts with the Pto kinase.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U89257; AAC49741.1; -
 DR PIR; T07728; T07728.
 DR HSSP; O80337; 2GCC.
 DR TRANSFAC; T04717; -
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PR00367; ETHRSPELEMT.
 DR PRODOM; PD001423; TF_ERF; 1.
 DR SMART; SM00380; AP2; 1.
 KW Transcription regulation; DNA-binding; Activator; Nuclear protein.
 FT DNA BIND 95 159
 FT SEQUENCE 248 AA; 27913 MW; 115BD44DDEB1B992 CRC64;
 SQ
 Query Match 4.6%; Score 145; DB 1; Length 248;
 Best Local Similarity 28.7%; Pred. No. 0.006;
 Matches 50; Conservative 31; Mismatches 67; Indels 26; Gaps 7;
 YQ 144 NQVDNVNDOE-NGNAAGLSLSMNSSTSCDNNNSNNVVAQGTIDDSVEATPKTIE 202
 DB 19 NKHFVSEHKEKNSLQVRVRIILTADATSSDDEGRTVRVRKHVTEINLMP--STK 76
 YQ 203 SFG--QRTSI-----YRGVTRHWRTGRYEAHLWDNSCKREGOTRGRQVILGY 249
 DB 77 SIGDRKRSVSPDSVTRRKPKRGVRQRPW-GRWAAEIRDPT-----RGRVWLGT 127
 YQ 250 DKEEKAARYDLAALKYWGTTTTTTFPMSEYEKEVEEMKHMTREQEYVASLRKS 303
 DB 128 DTPEEAAYVDKAAVKLKGPDATVTFPVST-TAEVTVTVTETETESVADGGDKS 180

```

RESULT 4
AS10 YEAST
ID AS10 YEAST STANDARD; PRT; 1146 AA.
AC P48361;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ASK10 protein.
GN ASK10 OR YGR097W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / SEY6210;
RX MEDLINE=97060018; PubMed=8904339;
RA Page N., Sheraton J., Brown J.L., Stewart R.S., Bussey H.;
RT "Identification of ASK10 as a multicopy activator of Skn7p-dependent
transcription of a HIS3 reporter gene.";
RL Yeast 12:267-272 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC Hernandez K., Weber N., Wipfli P., Schmidheini T.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Putative activator of SKN7.
CC -!- SIMILARITY: TO YEAST YIL105C AND YNL047C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U27209; AAA67368.1; -.
DR EMBL; U272882; CAA97100.1; -.
DR PIR; S64402; S64402.
DR GenOnline; 141409; -.
DR TRANSFAC; T03201; -.
DR SGD; S0003329; ASK10.
DR GO; GO:0006350; P:transcription; IGI.
DR InterPro; IPR001849; PH.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR DOMAIN 22; 26 POLY-GLY.
FT DOMAIN 625 628 POLY-SER.
FT DOMAIN 933 938 POLY-ASN.
FT DOMAIN 958 961 POLY-GLN.
FT DOMAIN 972 975 POLY-SER.
FT CONFLICT 57 57 I -> T (IN REF. 1).
FT CONFLICT 346 346 T -> N (IN REF. 1).
FT CONFLICT 464 464 P -> R (IN REF. 1).
FT CONFLICT 467 467 V -> A (IN REF. 1).
FT CONFLICT 603 603 A -> V (IN REF. 1).
FT CONFLICT 906 906 L -> P (IN REF. 1).
SQ SEQUENCE 1146 AA; 126863 MW; 9E9EF2B0BE496EE CRC64;

Query Match
Best Local Similarity 4.3%; Score 135.5; DB 1; Length 1146;
Matches 114; Conservative 89; Mismatches 187; Indels 217; Gaps 31;

QY 11 SPYEQNHRRKDVYSTTTTVDVAGCYDPTAASDESSAIOQTSFPSPFGVVDATRDN 70
DB 507 SNYIHEFOSSDFYLNSSS-----TPNSTKSSAYSSVS-----IADTVANAN 548
QY 71 NSHSDWDINGACNNHNDQDQKLEFLGRITTTIYNENVGDSGCGYGGGDDGG 130
DB 549 NAKA-----NNHHRQASD-----VHNSSTTT-----GGTAGAN 576

```

```

QY 131 GSLGLSNIKTWLRNPVDVNDQENGNAAGKGLSLSNSTSCDNNNDNNNVAAQK-TI 189
DB 577 GIRGIRK-KSYL-----APIMSIFLN-----DCTLDASSTKFLVVKPTL 616
QY 190 DDSVE-----ATPKTIESFGQRTS-IYRGVTRHRTWGRYEALHMDNSCKREGQ 237
DB 617 NENADVRSKSSSTYLSGSSQASLPKYGHETAKIPKAPFK----- 657
QY 238 TRKGRQVYLGGYDKEEKAARAYDL-----AALKYWGTTTTTNFPMSEYEKEVEEMKHWR 292
DB 658 -----FLKSKPKXNKTKSELDOFVAAQAKESNNVVTWTFKIVSPSPSEELKHKFR 710
QY 293 QEYVASLRKSKSFGS-----RGASIVRGVTRHGHQWQARIGRVAGNKDLVYG----- 341
DB 711 --VVQDLKNLTS-FNDTKDRIKFIEDRVNWKSHRF-----KAGHMSRN-SVNIHSHTPCLT 761
QY 342 --TFGTQEEAABAYDIAAIKFRGL-----TAVTNFDMRNYNVKAIKLESFSLPIG 388
DB 762 DSTFTLQDCTT-----TSVNLKGRAEKFCYIHIQNNSLADFDGNGFRSK--VNTFA-ID 812
QY 389 SAAKLKEANRPVPSMMISNNVSESENASCWQAAVQHGGVDLSLLHGHQBYNGY- 447
DB 813 DYGNLITVERRPAQSPHOQSDYMATSGNTTPSYSGS-----RQSMYNGYN 859
QY 448 -----YVINGGNLSSESARACFKQEDDQHHFASNTQSLMTNIDHSSVSDSDSVTCGVNVG 502
DB 860 PAVSITSGMMLQOSTA-----NNNTNPTTLRHRQHNISQTS-----SLPG 900
QY 503 YGYGQGFAPVNCDAAYASEFDYNARNHYVFAQQOOTOOSPGGPPAAMTNVNSNMYYH 562
DB 901 P-SYTSLSLPVNSPGSSNSE-----SSSGYFAIPLHGNNNNNTYTQ 941
QY 563 GEGGGEVAPFTVMDN 579
DB 942 RNSEGS-SPCY---NDD 954

RESULT 5
ID RAV1 ARATH STANDARD; PRT; 344 AA.
AC Q9ZNM9.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DNA-binding protein RAV1.
OS RAV1 OR ATIG13260 OR T6J4.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seedling;
RC MEDLINE=99081843; PubMed=9862967;
RA Kagaya Y., Ohmiya K., Hattori T.;
RT "RAV1, a novel DNA-binding protein, binds to bipartite recognition
RT sequence through two distinct DNA-binding domains uniquely found in
RT higher plants.";
RL Nucleic Acids Res. 27:470-478 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ecv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,

```

RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militscher J., Miranda M., Nguyen M., Niemman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W., of the plant Arabidopsis
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RL thaliana.";
RN Nature 408:816-820 (2000).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed14593172;
RA Yamada K., Lim J., Dale J.C., Kim C.J., Nguyen M., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Palm C.J., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.T., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjail M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RL Science 302:842-846 (2003).
CC -!- FUNCTION: Binds specifically to bipartite recognition sequences
CC composed of two unrelated motifs, 5'-CAACA-3' and 5'-CACCTG-3'.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in all tissues examined: Roots,
CC rosette leaves, cauline leaves, inflorescence stems, flowers and
CC siliques. Highest expression in roots and rosette leaves. Very low
CC expression in flowers.
CC -!- DOMAIN: Contains two distinct DNA-binding domains. One is located
CC in the N-terminal region and binds to the 5'-CAACA-3' motif. The
CC second is located in the C-terminal region and binds to the
CC 5'-CACCTG-3' motif.
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB013886; BAA34250.1; -;
DR EMBL; AC011810; AAG09554.1; -;
DR EMBL; AY063855; AAL36211.1; -;
DR EMBL; AY091291; AAM14230.1; -;
DR PIR; T51329; T51329.
DR HSSP; O80337; 2GCC.
DR InterPro; IPR003340; TF B3.
DR InterPro; IPR001471; TF ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR Pfam; PF02362; B3; 1.
DR PRINTS; PR00367; ETRSPLENT.
DR ProDom; PD001423; TF ERF; 1.
KW DNA-binding; Nuclear protein.
FT DNA BIND 59 121
FT DOMAIN 145 338
FT B3.
SQ SEQUENCE 344 AA; 38597 MW; 7349B640B3505823 CRC64;
Query Match 4.2%; Score 132.5; DB 1; Length 344;
Best Local Similarity 21.7%; Pred. No. 0.064;
Matches 68; Conservative 48; Mismatches 125; Indels 73; Gaps 10;
QY 177 DSNVNVVAGQKTTDDSVBATEPKKTIESPQRTSIYGVTRHRTWGRYEAHLWDNSCKREG 236

DB 45 DSENGVEAESKLP-----SSKYGVPQP-NGRWCAQIYE----- 79
QY 237 QTRKGRQYVLGYDKEEKAARAYDLAALKYWGTTTTTNF-PMSYEKEVEEMKMTQEQY 295
DB 80 ---KHORVWLGTFTNEDEEAAAYDVAVHRFRFRDAVTFKDVKMDEDEVDPLNSHSEI 136
QY 236 VASLRKSSGFSRGASIVGYTRHHQGRWQCARIGRVAGNKDLVLTGFTGTEAAEAYDI 355
DB 137 VMLRKHTYNEELSQS-----KRRNGN-----GNTRTLTSLGSLNDGVSTTGF 181
QY 356 AAIKFRGLTAVTNFDMRNYNKAI-----LESFSLPIGSAARLKEANRPVPSMM 405
DB 182 RSAEALFEKAVTPSDVGVKLNELVLPKHAERKHFPLPSSNVSVKGVLLNFEDVNGKWRFR 241
QY 406 MISNVSSSENSASQWQAQVHH-----QGVDSLHLH-QHQRINYGYNYNG- 452
DB 242 YSYWNSSSYVLTGKWSRFVKEKNLRAGDVVYFSRSGNQDQOLYIGWKSRSGLDAGR 301
QY 453 -----NLSSESAR 460
DB 302 LRLFGVNISSPSSR 315
RESULT 6
ID ERF2 ARATH STANDARD; PRT; 243 AA.
AC O80338;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ethylene responsive element binding factor 2 (AtERF2).
GN ERF2 OR ERF-2 OR AT5G47220 OR MQL5.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181733; PubMed=10715325;
RA Fujimoto S.Y., Ohta M., Usui A., Shinshi H., Ohme-Takagi M.;
RT "Arabidopsis ethylene responsive element binding factors act as
RT transcriptional activators or repressors of GCC box mediated gene
RT expression.";
RL Plant Cell 12:393-404 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RL clones";
RL DNA Res. 7:31-63 (2000).
CC -!- FUNCTION: Transcription factor that binds to the GCC-box
CC pathogenesis-related promoter element.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB008104; BAA32419.1; -;
DR EMBL; AB018117; BAA97155.1; -;
DR PIR; T51989; T51989.
DR HSSP; O80337; 2GCC.


```
FT PROPEP ? 1291 POTENTIAL.
SQ SEQUENCE 1291 AA; 139635 MW; ECA56A61CAE36669 CRC64;

Query Match
  4.1%; Score 126.5; DB 1; Length 1291;
Best Local Similarity
  18.0%; Pred. No. 0.97;
Matches 11; Conservative 73; Mismatches 191; Indels 241; Gaps 27;

QY 1 MNNNL-----GFLSPVEQNHRRKDVSSSTTTVVVAGEVCYDPTAASDESSAIOIS 54
DB 254 MGNVWGRQYQVAYLAP-----SYSTINTSKVTGEVFNHLTVGDHNA----- 298
QY 55 FSPFGVGVDAFTRDNNSHSRDWDI-----NGACNINHND 90
DB 299 ---QAGIIA-----SNKTHITGLDLQWSAGLNIAPPBGYKDKPKDNTSTNNAKND 350
QY 91 EQ-----DQPLENLFGRITTTTYNNENNVGDSGSG 121
DB 351 KQSSQNNSTQVNPNSAQTETIQTOVIDP-----FAGGKNTVVNIRNTNADGII 406
QY 122 CYGGDGB-----GGSLGLS-----MKTWLRNOPVD-----NVDNQNENNA 159
DB 407 RVGGFKASLTNAAHLHIGKGINLSQASGRSLVENLTGNTVDGPLRVNNGVGYAL 466
QY 160 KGLSLNNSSTSDNNNSN-----NNVWAQKTIIDSVEATEPKTIESFGORTSYRGVTRH 217
DB 467 AGSSANPEFKAGDTKNGTATFNNDISLGRFVNLKVD-----HTANFKGID----- 513
QY 218 RWTGRYEHLWDNSCKREGQTRKGRQVYLGQYDKEEKAARAYDLAALKYWGTTTTTFPM 277
DB 514 --TGN-----GGFN-----TLDFSGVTNKVNINK 535
QY 278 SEYEKEVEEMKHTROFYVASLRKSGSFGSASIVRGVTRHHQHGWSQARIGRV---AG 334
DB 536 LITASTNVAVKNFENELV-----KINGVSVE-----YTHFSEDIGSQSRINTVLETG 586
QY 335 NKDLYLTGFTQBEAAEYDIAIKFRLTAFTNFD-----MNRYNVKALESFSLP 386
DB 587 TRSYSG-----GVKPKGKLVINDFYAPVYFADARNIKNVEINKLA 631
QY 387 IGSAAKRLKANRPVPMMSINNVSENSA---SGWQAAVQ---HHQGVDSLHLHQ 439
DB 632 FGQGS-----PWGTAKLFMNLTLGQNAVMDYSQFSNLTIGQDFVNNQGT----- 677
QY 440 HQERYNGYNYNGG-----NLSSSARACFKQEDDQHHF-----LSNTQSLMTNIDHQS 488
DB 678 -----INLYVGGQVATLVGNAAAMFFSNVDSATGFGYQPLMKINSAQDLINKKEH--- 729
QY 489 VSDSVTVCGNVVGYG 504
DB 730 -----VLKAKIICYG 740

RESULT 9
ERF4 ARATH STANDARD; PRT; 222 AA.
ID ERF4 ARATH
AC O80340;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ethylene responsive element binding factor 4 (ATERF4).
GN ERF4 OR ERF-4 OR AT3G15210 OR K7L4.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181733; PubMed=10715325;
RA Fujimoto S.Y., Ohta M., Usui A., Shirshi H., Ohme-Takagi M.;
RT "Arabidopsis ethylene responsive element binding factors act as
transcriptional activators or repressors of GCC box mediated gene
expression."
```

Plant Cell 12:393-404(2000).

RL RN SEQUENCE FROM N.A.

RP STRAIN=sv Columbia;

RX MEDLINE=21016720; PubMed=11130713;

RA Salanoubat M., Lemcke K., Rieger H., Perez-Alonso M., Oberwälder B.,

RA Farmanan B., Valle G., Bloeker M., Maché R., Puigdomenech P.,

RA Delany M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,

RA De Simone V., Choisine N., Artiguenave F., Robert C., Brottier P.,

RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quétier F.,

RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,

RA Wurmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,

RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,

RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,

RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,

RA Reichelt J., Schärfe M., Schoen O., Barges M., Terol J., Climent J.,

RA Navarro P., Collado C., Perez-Perez A., Ottenwälder B., Duchemin D.,

RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,

RA de Haan M., Maere A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,

RA Monfort A., Argüello A., Flores M., Liguori R., Vitale D.,

RA Manthaupt G., Haase D., Schoof H., Rued S., Zaccaria P., Mewes H.-W.,

RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,

RA Rooney T., Rizzo M., Waits A., Utterback T., Fujii C.Y., Shea T.P.,

RA Creasy T.H., Haas B., Waiiti R., Wu D., Peterson J., Van Aken S.,

RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,

RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,

RA Fraser C.M., Kimura T., Nakamura Y., Sato S., Kato T., Asamizu E.,

RA Sasamoto S., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

RA Kiyokawa C., Nakazaki N., Shimpō S., Takeuchi C., Wada T.,

RA Nakayama S., Yamada M., Yasuda M., Tabata S.,

RA Watanabe A., Yamada M., Yasuda M., Tabata S.,

RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis

RL Nature 408:820-822(2000).

CC -!- FUNCTION: Transcription factor that binds to the GCC-box

CC pathogenesis-related promoter element.

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -!- SIMILARITY: Contains 1 AP2/ERF domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; AB008106; BAA32421.1; -;

DR EMBL; AF000413; BAB02150.1; -;

DR PIR; T52019; T52019.

DR HSP; O80337; 2GCC.

DR TRANSFAC; T04636; -;

DR InterPro; IPR001471; TF_ERF.

DR Pfam; PF00847; AP2-domain; 1.

DR PRINTS; PR00367; ETHRSEPELMT.

DR ProDom; PD001423; TF_ERF; 1.

DR SMART; SM00380; AP2; 1.

KW Transcription regulation; DNA-binding; Nuclear protein.

FT DNA BIND 22 86 AP2/ERF.

FT DOMAIN 132 137 POLY-GLY.

SQ SEQUENCE 222 AA; 23741 MW; CEF16B8A5ACFB073 CRC64;

Query Match 4.0%; Score 125; DB 1; Length 222;

Best Local Similarity 33.6%; Pred. No. 0.11;

Matches 39; Conservative 12; Mismatches 39; Indels 26; Gaps 5;

QY 197 PKKTIESFQRTSI-YRGVTRHRTGRYEHLWDNSCKREGQTRGROVYLGVDKEKA 255

DB 10 PATFTNQTHNAKEIRYGRKEPW-GRVAAEIRDPG-----KKTRVWLTFDTAEHA 60

QY 256 ARAYDLAALKYWGTTTTTTTFPMSEYEKEVEEMKHTROFYVASLRKSGSFGSRGAS 311

DB 61 ARAYDTARDPRGAKATNFPTF---LELSDQKVPT-----GPARSPS 100

RESULT 10
 ERFL1 ARATH
 ID ERFL1 ARATH STANDARD; PRT; 268 AA.
 AC ERFL1 ARATH STANDARD; PRT; 268 AA.
 DT 28-FEB-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ethylene responsive element binding factor 1 (ATERF1) (ZREBP-2 protein)
 GN ERFL1 OR ERF-1 OR AT4G17500 OR DL4785W.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TAXID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20181733; PubMed=10715325;
 RA Fujimoto S.Y., Ohta M., Usui A., Shinshi H., Ohme-Takagi M.;
 RT "Arabidopsis ethylene responsive element binding factors act as transcriptional activators or repressors of GCC box mediated gene expression.";
 RT Plant Cell 12:393-404 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98121113; PubMed=9461215;
 RA Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C.,
 RA Bervankamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
 RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P.,
 RA Wedler H., Wedler E., Wambutt R., Weitzenecker T., Pohl T., Terryn N.,
 RA Gielen J., Villarroel R., De Clercq R., van Montagu M., Lecharny A.,
 RA Aubourg S., Gil I., Kreis M., Lao N., Kavanagh T., Hempel S.,
 RA Kottler P., Enlian K.-D., Rieger M., Schaefer M., Funk B.,
 RA Mueller-Auer S., Slivey M., James R., Monfort A., Pons A.,
 RA Puigdomenech P., Douka A., Vouklatou E., Milioni D., Hatzopoulos P.,
 RA Piravandi E., Obermaier B., Hilbert H., Duesterhoef A., Moores T.,
 RA Jones J.D.G., Enava T., Palme K., Benes V., Rechmann S., Ansoorge W.,
 RA Cooke R., Berger C., Delseny M., Voet M., Volckaert G., Mewes H.-W.,
 RA Klosterman S., Schueller C., Chalvatzis N.;
 RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana.";
 RT Nature 391:485-488 (1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoef A., Stiekema W., Enlian K.-D., Terryn N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grynoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., Mcay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borxova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., Fuhs M., Martmann B., Schaefer M., Mueller-Auer S.,
 RA Gabel C., De Haan M., Faarman A.C., Schaefer M., Dauner D., Herzl A.,
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley P., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Facon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bieleke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stonking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dente M., Pepin K., Hillier L.W.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.";
 RT Nature 402:769-777 (1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Haysashiaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema B., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis genome.";
 RT Science 302:842-846 (2003).
 RN [5]
 RP STRUCTURE BY NMR OF 146-208
 RX MEDLINE=98409552; PubMed=9736626;
 RA Allen M.D., Yamasaki K., Ohme-Takagi M., Tateno M., Suzuki M.;
 RT "A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA.";
 RT EMBO J. 17:5484-5496 (1998).
 RN [6]
 RP INDUCTION.
 RX MEDLINE=99069218; PubMed=9851977;
 RA Solano R., Stepanova A.N., Chao Q., Ecker J.R.;
 RT "Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1.";
 RT Genes Dev. 12:3703-3714 (1998).
 CC -!- FUNCTION: Transcription factor that binds to the GCC-box pathogenesis-related promoter element.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- INDUCTION: Ethylene induction of ERF1 is completely dependent on a functional ETHYLENE-INSENSITIVE3 (EIN3) protein.
 CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
 CC -!- CAUTION: Ref.2 and Ref.3 sequences differ from that shown due to a frameshift in position 44.
 CC -!- CAUTION: It is uncertain whether Met-1 or Met-3 is the initiator.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dentin sialophosphoprotein precursor [Contains: Dentin phosphoprotein
 DE (Dentin phosphophoryn) (DPP); Dentin sialoprotein (DSP)].
 GN DSPP.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20168992; PubMed=10706475;
 RA Gu K., Chang S.R., Ritchie H.H., Clarkson B.H., Rutherford R.B.;
 RT "Molecular cloning of a human dentin sialophosphoprotein gene.";
 RL Eur. J. Oral Sci. 108:35-42(2000).
 RN [2]
 RN SEQUENCE OF 463-1253 FROM N.A.
 RC TISSUE=Tooth;
 RX MEDLINE=99094526; PubMed=9879917;
 RA Gu K., Chang S.R., Slaven M.S., Clarkson B.H., Rutherford R.B.,
 RA Ritchie H.H.;
 RT "Human dentin phosphophoryn nucleotide and amino acid sequence.";
 RL Eur. J. Oral Sci. 106:1043-1047(1998).
 RN [3]
 RN DISEASE.
 RX MEDLINE=21096971; PubMed=11175779;
 RA Zhang X., Zhao J., Li C., Gao S., Qiu C., Liu P., Wu G., Qiang B.,
 RA Lo W.H.Y., Shen Y.;
 RT "DSP mutation in dentinogenesis imperfecta Shields type II.";
 RL Nat. Genet. 27:151-152(2001).
 RN [4]
 RN VARIANTS DFNA39/DGII THR-17 AND PHE-18.
 RX MEDLINE=21096982; PubMed=11175790;
 RA Xiao S., Yu C., Chou X., Yuan W., Wang Y., Bu L., Fu G., Qian M.,
 RA Yang J., Shi Y., Hu L., Han B., Wang Z., Huang W., Liu J., Chen Z.,
 RA Zhao G., Kong X.;
 RT "Dentinogenesis imperfecta 1 with or without progressive hearing loss
 RT is associated with distinct mutations in DSPP.";
 RL Nat. Genet. 27:201-204(2001).
 CC !- FUNCTION: DSP may be an important factor in dentinogenesis. DPP
 CC may bind high amount of calcium and facilitate initial
 CC mineralization of dentin matrix collagen as well as regulate the
 CC size and shape of the crystals.
 CC !- SUBCELLULAR LOCATION: Secreted.
 CC !- TISSUE SPECIFICITY: Expressed in teeth. DPP is synthesized by
 CC odontoblast and transiently expressed by pre-ameloblasts.
 CC !- PTM: DSP is glycosylated.
 CC !- DISEASE: Defects in DSP are the cause of dentinogenesis
 CC imperfecta 1 (DGII) (MIM:125490); also known as dentinogenesis
 CC imperfecta Shields type II (DGII-II). DGII is an autosomal dominant
 CC disorder in which both the primary and the permanent teeth are
 CC affected. It occurs with an incidence of 1:8000 live births. The
 CC teeth are amber and opalescent, the pulp chamber being obliterated
 CC by abnormal dentin. The enamel, although unaffected, tends to
 CC fracture, which makes dentin undergo rapid attrition, leading to
 CC shortening of the teeth.
 CC !- DISEASE: Defects in DSP are the cause of autosomal dominant
 CC deafness type 39 with dentinogenesis imperfecta 1 syndrome
 CC (DFNA39/DGII) (MIM:605594). Affected individuals present DGII
 CC associated with early onset progressive sensorineural high-
 CC frequency hearing loss.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF163151; AAF42472.1; -;
 DR EMBL; AF094508; AAD16120.1; -;

DR Genew; HGNC:3054; DSPP.
 DR MIM; 125485; -;
 DR MIM; 125490; -;
 DR MIM; 605594; -;
 KW Biomineralization; Extracellular matrix; Signal; Glycoprotein;
 KW Phosphorylation; Sialic acid; Disease mutation; Deafness.
 FT SIGNAL 1 15
 FT CHAIN 16 1253
 FT CHAIN 16 462
 FT CHAIN 463 1253
 FT DOMAIN 439 1253
 FT SITE 488 490
 FT MOD_RES 259 259
 FT CARBOHYD 41 41
 FT CARBOHYD 43 49
 FT CARBOHYD 81 81
 FT CARBOHYD 130 130
 FT CARBOHYD 150 150
 FT CARBOHYD 190 190
 FT CARBOHYD 191 191
 FT CARBOHYD 209 209
 FT CARBOHYD 222 222
 FT CARBOHYD 275 275
 FT CARBOHYD 336 336
 FT CARBOHYD 387 387
 FT VARIANT 17 17
 FT VARIANT 18 18
 FT VARIANT 18 18
 FT CONFLICT 685 690
 FT CONFLICT 744 744
 FT CONFLICT 799 799
 FT CONFLICT 836 836
 FT CONFLICT 850 850
 FT CONFLICT 886 888
 FT CONFLICT 963 963
 FT CONFLICT 1005 1005
 FT CONFLICT 1025 1025
 FT CONFLICT 1047 1047
 FT CONFLICT 1053 1053
 FT CONFLICT 1065 1065
 FT CONFLICT 1080 1080
 FT CONFLICT 1095 1095
 FT CONFLICT 1101 1101
 FT CONFLICT 1104 1104
 FT CONFLICT 1132 1132
 FT CONFLICT 1253 1253
 SQ SEQUENCE 1253 AA; 126426 MW; DAE240653904ED4A CRC64;
 Query Match 3.9%; Score 122.5; DB 1; Length 1253;
 Best Local Similarity 18.7%; Pred. No. 1.7; Indels 141; Gaps 21;
 Matches 102; Conservative 80; Mismatches 222;
 QY 43 AASDESSAIQTSPFPFGVYVDAFTDRNNSHSRWDINGCACNNIHNDEQPKLENFLG 102
 DB 43 SYQDELNASGTIKES--GVLVHEGDRGRQENTQDG-----HKGEHGSKWAEVGG 90
 QY 103 RYTTNTNTNENVDGSGGCGGCGGGLSLMIKTLWRNPQVNDVNDQNGNAAGKL 162
 DB 91 KGFSTYSTLAN--EGNIEG--WNGDTGKAETYG-----HDGIGHKEENTANGI 136
 QY 163 ----SLSMNSSTSCDNNNDNNVVAQGTITDVSVEATPKKTIIESFGORTSIYRGVTRHR 218
 DB 137 QGVSIIDNAGATNRSNTNGNTDKNTQNGDVGADHNEVDVAVQEDGPQVAGSNNSTDN- 195
 QY 219 WTGRVEAHUWNSCKRGGOT-----RKGRVYLGVDYKKEKAAARAYDLAAKTKWG 268
 DB 196 -----EDEIENSRCNEGNTSITPQINSKNGTK-----EAEVTPGTGDA-----G 238
 QY 269 TTTTTFNPFMSEYKEVEEMKMTROYVVASLRKKSFGSRGASIVRGVTRHHQGRWQAR 328
 DB 239 LNSDGSFSGNGADEDEDEGGCGDDEEAGNGKDSNNNSKQE----- 281
 QY 329 IGRVAGNKDLYLGTGTGTQTEFAAEADIAAIKFRGLTAVTNFDMNRYNKAILESPSLPIG 388

RP SEQUENCE FROM N.A.
 RC STRAIN=CCUG 17874 / NCTC 11638;
 RX MEDLINE=94222514; PubMed=8168917;
 RA Phadnis S.H., Ilver D.J., Janson L., Normark S., Westblom T.U.;
 RT "Pathological significance and molecular characterization of the
 RL vacuolating toxin gene of *Helicobacter pylori*";
 RN Infect. Immun. 62:1557-1565(1994).
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RA MEDLINE=94216833; PubMed=8163943;
 RX Telford J.L., Ghara P., Dell'Orco M., Comanducci M., Burroni D.,
 RA Bugnoli M., Tecce M.P., Censini S., Covacci A., Xiang Z., Papini E.,
 RA Montecucco C., Parente L., Rappuoli R.;
 RT "Gene structure of the *Helicobacter pylori* cytotoxin and evidence of
 RL its key role in gastric disease";
 RN J. Exp. Med. 179:1653-1658(1994).
 CC -!- FUNCTION: Induces vacuolation of eukaryotic cells. Causes
 CC ulceration and gastric lesions.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U07145; AAA18867.1; -
 DR EMBL; S72494; AAB30582.1; -
 DR InterPro; IPR006315; Autotransporter.
 DR InterPro; IPR005546; Autotransporter.
 DR InterPro; IPR003842; Vaca.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02891; Vaca; 1.
 DR PRINTS; PRO1656; VACCYTOTOXIN.
 DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
 KW Toxin; Signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 ? VACUOLATING CYTOTOXIN.
 FT PROPEP 2 1296 ? POTENTIAL.
 FT PROPEP 2 1296 ? POTENTIAL.
 SQ SEQUENCE 1296 AA; 139760 MW; 0D1F3F71AB411447 CRC64;
 Query Match 3.9%; Score 121.5; DB 1; Length 1296;
 Best Local Similarity 18.8%; Pred. No. 2.1;
 Matches 138; Conservative 83; Mismatches 224; Indels 289; Gaps 39;
 QY 1 MNNWNL-----GFSLSPEYQNHKKDYVSSTTTTVDVAGEYCYDPTAASDESSAIGTS 54
 DB 254 MGNVWAGSLQVGVAYLAP-----SYSTINTSKVTGEVNFNHLTVGDKNA----- 298
 QY 55 FPSPPGVVDVDAFT-----RDNNSHRDWM-----DINGCANN 86
 DB 299 ---QAGIIANKKNTIGTLQWSAGLNTIAPPEGYKDKPNPTSQSGAKNDKVESAKND 355
 QY 87 IHNDEQ-----DGPKEFLGRTTITNTYNTNENVDGSGSG 121
 DB 356 KOESSQNSNTQVNPNSAQTVEQPTQVTDGP-----FAGGKDTVNNIRNTNADGTI 411
 QY 122 CVGG-----GDGG-----GGSLGSLMIKTLWRNPQVD---NVDNCENG 156
 DB 412 RVGGFKASLTNAHLHGKGVNLSNQAASGSL---IVENLTGNTIIVDGLRNVNQG 468
 QY 157 NAAKGLSLSMNSSTSCDNNDSN--NNVAGKVIDDSVEATPKKTTESFGQRTSIYRGV 214
 DB 469 YALAGSSANFBFKAGTDTKNGTATFNNDISLGRFVNLKQDA-----HTANFKGI 517
 QY 215 TRHRWTGRYEHLMDNSCKREGQTKRGVQLVGGYDKEEKAARAYDLAALKYWGTTTTTN 274
 DB 518 D-----TGN-----GGFN-----TLDFSVTDKVN 537
 QY 275 FPMSEYEKEVEMKHTMTQEVYASLRKSSGFSRGASIIYRGVTRHHQHRWQARIGV-- 332

Db 538 INKLITASTNVAQNFNINELIV---KTNGISVGE-----YTHFSEDIGSOSRINTVRL 588
 QY 333 -AGNKDLYLTGTGTGTOBEAAEAYDIAIKFRG-----LTAVTNFDMRYNNVKAILE 381
 Db 589 ETGTRSLFSG-----GVKFKGKELVIDEFYSPWNYFDAR--NIKWEI 631
 QY 382 SPSLPISGSAKLKEANRPVPMNMISNNVSENSA---SQWQAAYQ-----HHCQVDL 434
 Db 632 TNKLAFPGQS-----PWGTSKLMFNLTGLQAVMDYSQFSLNLTIOGDFINNQGT-- 682
 QY 435 SLLHQERYNGVYVNGGNLSS-----ESARACFKOE-DDQHHF-----LSNTQSLMTNI 483
 Db 683 -----INLVRGKGVATLSVGNAAAMFNNDIDSATGYFKPLIKINSADLIKNT 732
 QY 484 DQSSVSDSVTVCGNVVYG-----GYOGFAAPVNCDAVASEFD--YNARNHYYPAAQCCQ 538
 Db 733 EH-----VLLKAKIIGYGVNSTGTNGI--SNVNLEEQFKERLALYNNNRMDTCCVVRN 783
 QY 539 TQSPG-----GDFPAAATNN-----VGS-----NNVYHG-----EG 565
 Db 784 TDDIRACGMAIGD--QSVNVPDNYKLLGKAWNIGISKTAGSKISVYVLGNSTPTEN 841
 QY 566 GGEVA--PTFTVMN 577
 Db 842 GGNNTNLPTNT--SN 855
 RESULT 15
 ID GPRI_YEAST STANDARD; PRT; 961 AA.
 AC Q12361;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE G protein-coupled receptor GPRI.
 GN GPRI OR YDL035C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98190081; PubMed=9524122;
 RA Xue Y., Batlle M., Hirsch J.P.;
 RT "GPRI encodes a putative G protein-coupled receptor that associates
 RT with the Gpa2p Galpha subunit and functions in a Ras-independent
 RT pathway";
 RL EMBO J. 17:1996-2007(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=97197972; PubMed=9046088;
 RA Saren A.M., Laamanen P., Lejarcegui J.B., Paulin L.;
 RT "The sequence of a 36.7 kb segment on the left arm of chromosome IV
 RT from *Saccharomyces cerevisiae* reveals 20 non-overlapping open reading
 RT frames (ORFs) including SIT4, FAD1, NAM1, RNA11, SIR2, NAT1, PRP9, ACT2
 RT and MFS1 and 11 new ORFs";
 RL Yeast 13:65-71(1997).
 CC -!- FUNCTION: Seems to associate with GPA2 and act as G protein-
 CC coupled receptor that senses glucose and controls filamentous
 CC growth. It acts upstream of adenylate cyclase and is required for
 CC glucose activation of cAMP synthesis in concert with a glucose
 CC phosphorylation-dependent mechanism.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- SIMILARITY: BELONGS TO FAMILY GPRI/GIT3 OF G-PROTEIN COUPLED
 CC RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z74083; CAA98593.1; -;
DR EMBL; Z71781; CAA96454.1; -;
DR PIR; S67568; S67568.
DR Germonline; 140277; -;
DR SGD; S0002193; GPR1.
DR GO; GO:0005886; Cytoplasm membrane; IDA.
DR GO; GO:0004930; F-G-protein coupled receptor activity; IMP.
DR GO; GO:0007186; F-G-protein coupled receptor protein signalin. .; IMP.
DR GO; GO:0007124; Pseudohyphal growth; IMP.
KW G-protein coupled receptor; Transmembrane.
FT TRANSMEM 57 79
FT TRANSMEM 92 114 POTENTIAL.
FT TRANSMEM 134 156 POTENTIAL.
FT TRANSMEM 179 198 POTENTIAL.
FT TRANSMEM 251 273 POTENTIAL.
FT TRANSMEM 620 642 POTENTIAL.
FT DOMAIN 501 547 POLY-ASN.
SQ SEQUENCE 961 AA; 110708 MW; 9889D85782A4209 CRC64;

Query Match 3.98; Score 121; DB 1; Length 961;
Best Local Similarity 18.0%; Pred.No.1.5;
Matches 105; Conservative 83; Mismatches 201; Indels 194; Gaps 28;

Qy 9 SLSPYE----QNHRKDVYSGSTTTTVDVAGCYCDPTAASDESSAIQTSPFPFGVVD 64
Db 420 TLNPSLVSKQEHK-----VTFSVENEGLDTRKSSMLGHQTFSCQNSLESPLAVY-- 470

Qy 65 APTRDNNSHRDWDIN-----GCACNNIHNDQGGPKLENFLGRTTTIYNTNENVGDSG 119
Db 471 ----DNKNDNSDITSNIKKEGGIINNNSNDDDDN-----NNNDNDNDNN 513

Qy 120 SGCVGGDGGGSLGLSMKTLNRPQVDVNDQENGNAAGLSLWSNSTSCDNNDSN 179
Db 514 S-----NNNNNNNNNN-----NNNNNNNNNN 541

Qy 180 NNVAQOKTIDDSVEATPKXTI-----ESFGQ-----RTSIYRGVTRHR 218
Db 542 NNSNNIKNNVDNNTNPADNPTLSNEAFTPSQSFQSERVNNNADRCENSFTNVQH- 600

Qy 219 WTGRYEHLWDNSCKREGQTRKG-RQVYLGVDYDKEEKAARAYDLAALKYW----- 267
Db 601 ----FQAQTYQMKRRRAQIQKNLRAIFI-----YPLSYIGIWLFIADALQ 644

Qy 268 -----GTTTTTNF-----PMS-----EYKEVEEMKHMTRQYVAS- 298
Db 645 YNHEIKHGPTMWVTYIDTCVRLSCLVDVIVLPKRPWYNSWAKTESKYLI-EKYLKG 703

Qy 299 -----LRRKSSGFSRGSIYRGVTRHHQGRWOARIGRVAGNKDLYLGTGTQBEAAE 351
Db 704 ELGEKEILKFTCHSNWGRGWYR-----GKWKKEKCKWYKSTNPLKRLWFWVERFFKQ 755

Qy 352 AYDIAAIKFRGLTAVTNFDM--NRYNVKAILSPSPISAAKRLKEANRPVPS-----M 404
Db 756 LFEL-KLHFSFYDNCDDFEYWNYSYAKD--SNDNKESTEDTKNTSSDRSLPSNLELQ 812

Qy 405 MWTISNVSENSASGQNAVQHH-----QGVDSLHLHQCHERYNGYINGNLSSESAR 460
Db 813 AMLNNITAAEEVEVPLFWR---IIHHIPLMGIDLDLNLN----RLIKIRYNNHDFSLPGLK 865

Qy 461 ACFKQ--EDQOHFLSN-----TQSLMTNIDHQSSVDD 492
Db 866 PALNQKSHDKHQDVSTNSMVKSFSSNIVTN--DDNSIEED 907

Search completed: July 15, 2004, 10:02:16
Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 09:58:34 ; Search time 43 Seconds
(without alignments)
4248.489 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119

Sequence: 1 MNNNLGFSLSPEYQNHKK.....YVHGEGGEVAFPTFTWINDN 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_protein.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriophage.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3119	100.0	579	10 Q8L3U3	Q8L3U3 brassica na
2	3041	97.5	579	10 Q8LSN2	Q8LSN2 brassica na
3	2484.5	79.7	584	10 Q8LSN1	Q8LSN1 arabidopsis
4	2475	79.4	581	10 Q9LFP42	Q9LFP42 arabidopsis
5	1139	36.5	655	10 Q7X6D4	Q7X6D4 oryza sativ
6	1131	36.3	692	10 Q8S1E1	Q8S1E1 oryza sativ
7	1089.5	34.9	597	10 Q8LGO3	Q8LGO3 oryza sativ
8	1040.5	33.4	516	10 Q9SYC2	Q9SYC2 arabidopsis
9	997.5	32.0	540	10 Q9LTX37	Q9LTX37 arabidopsis
10	997	32.0	566	10 Q9LXA3	Q9LXA3 arabidopsis
11	971.5	31.1	639	10 Q84Z02	Q84Z02 oryza sativ
12	963	30.9	492	10 Q8LGO4	Q8LGO4 oryza sativ
13	950.5	30.5	555	10 Q9FIE2	Q9FIE2 arabidopsis
14	930	29.8	485	10 Q41832	Q41832 zea mays (m
15	926	29.7	425	10 Q9CAH3	Q9CAH3 arabidopsis
16	922.5	29.6	437	10 Q9LSM4	Q9LSM4 arabidopsis

ALIGNMENTS

RESULT 1

Q8L3U3 PRELIMINARY; PRT; 579 AA.
 ID Q8L3U3
 AC Q8L3U3;
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE AP2/EREP transcription factor BABY BOOM1.
 GN BM1.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]_TaxID=3708;
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Topas;
 RA Boutiller K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,
 RA Quillet T., Zhang L., Hattori J., Liu C.-M., Miki B.L.A.,
 RA Custers J.B.M., van Lookeren Campagne M.M.;
 RT "Ectopic expression of the Brassica napus BABY BOOM gene triggers a
 conversion from vegetative to embryonic growth.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317904; AAM33800.1; -;
 DR EMBL; AF317906; AAM33802.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 2.
 DR PRINTS; PR00367; ETRSPLEMT.
 DR ProDom; PD001423; TF_ERF; 2.
 DR SMART; SM00380; AP2; 2.
 SQ SEQUENCE 579 AA; 63962 MW; EA2APB4734500D72 CRC64;

Query Match 100.0%; Score 3119; DB 10; Length 579;

Best Local Similarity 100.0%; Pred. No. 8.6e-204;

Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNWLGFSLSPEQNHHRKDYVSSTTTTVDVAGEYCYDPTAASDBSSAIQTSPSPFG 60
 Db 1 MNNWLGFSLSPEQNHHRKDYVSSTTTTVDVAGEYCYDPTAASDBSSAIQTSPSPFG 60
 QY 61 VVVDFTTRNNSHSRDWDINGCACNNIHNDEODGPKLENFLGRITTTIYNTNENVDGSGS 120
 Db 61 VVVDFTTRNNSHSRDWDINGCACNNIHNDEODGPKLENFLGRITTTIYNTNENVDGSGS 120
 QY 121 GYGGGDDGGGSLGSLSMIKTWLRNQPVDNVDNQENGAAGKLSLSMNSSTSCDNNNSNN 180
 Db 121 GYGGGDDGGGSLGSLSMIKTWLRNQPVDNVDNQENGAAGKLSLSMNSSTSCDNNNSNN 180
 QY 181 NVVAQKTIIDSVATPKKTIIESFGQRTSIYRGVTRHRTWGRYEHLWDNSCKREGQTRK 240
 Db 181 NVVAQKTIIDSVATPKKTIIESFGQRTSIYRGVTRHRTWGRYEHLWDNSCKREGQTRK 240
 QY 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTTPMSEYEKEVEEMKHMTRQYVASLR 300
 Db 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTTPMSEYEKEVEEMKHMTRQYVASLR 300
 QY 301 RKSSGFSRGASIRYGVTRHHQGRWQARIGRVAGNKDLYLGTFTGQEEAAAYDIAAIKF 360
 Db 301 RKSSGFSRGASIRYGVTRHHQGRWQARIGRVAGNKDLYLGTFTGQEEAAAYDIAAIKF 360
 QY 361 RGLTAVTNFDMNRNVKAILSPISLPISGAARLKEANRPVPSMMTISNNVSESENSASG 420
 Db 361 RGLTAVTNFDMNRNVKAILSPISLPISGAARLKEANRPVPSMMTISNNVSESENSASG 420
 QY 421 WNAVAQHOGVDLSLLHQHQRNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480
 Db 421 WNAVAQHOGVDLSLLHQHQRNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480
 QY 481 TNIDHQSSVSDSVTCGVNMGYGGYQGFAPVNCDAVAASEFDYNARNHYFFAQOQQTQ 540
 Db 481 TNIDHQSSVSDSVTCGVNMGYGGYQGFAPVNCDAVAASEFDYNARNHYFFAQOQQTQ 540

RESULT 2

Q8LSN2 PRELIMINARY; PRT; 579 AA.
 ID Q8LSN2
 AC Q8LSN2
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE AP2/ERBP transcription factor BABY BOOM2.
 GN BBM2.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ev. Topas.
 RA Boutiller K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,
 RA Zhang L., Hattori J., Ouellet T., Liu C.-M., Miki B.L.A.,
 RA Custers J.B.M., van Lookeren Campagne M.M.;
 RT "Ectopic expression of the Brassica napus BABY BOOM gene triggers a
 RT conversion from vegetative to embryonic growth.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317905; AAM33801.1;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR01471; TF_ERF.
 DR Pfam: PF00847; AP2-domain; 2.
 DR PRINTS: PR00367; ETRSRLEMT.
 DR ProDom: PD001423; TF_ERF; 2.
 DR SMART: SM00380; AP2; 2.

SQ SEQUENCE 579 AA; 64018 MW; 6426D3B2B06D3BEA CRC64;
 Query Match 97.5%; Score 3041; DB 10; Length 579;
 Best Local Similarity 97.6%; Pred. No. 1.7e-198;
 Matches 565; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
 QY 1 MNNWLGFSLSPEQNHHRKDYVSSTTTTVDVAGEYCYDPTAASDBSSAIQTSPSPFG 60
 Db 1 MNNWLGFSLSPEQNHHRKDYVSSTTTTVDVAGEYCYDPTAASDBSSAIQTSPSPFG 60
 QY 61 VVVDFTTRNNSHSRDWDINGCACNNIHNDEODGPKLENFLGRITTTIYNTNENVDGSGS 120
 Db 61 VVVDFTTRNNSHSRDWDINGCACNNIHNDEODGPKLENFLGRITTTIYNTNENVDGSGS 120
 QY 121 GYGGGDDGGGSLGSLSMIKTWLRNQPVDNVDNQENGAAGKLSLSMNSSTSCDNNNSNN 180
 Db 121 GYGGGDDGGGSLGSLSMIKTWLRNQPVDNVDNQENGAAGKLSLSMNSSTSCDNNNSNN 180
 QY 181 NVVAQKTIIDSVATPKKTIIESFGQRTSIYRGVTRHRTWGRYEHLWDNSCKREGQTRK 240
 Db 181 NVVAQKTIIDSVATPKKTIIESFGQRTSIYRGVTRHRTWGRYEHLWDNSCKREGQTRK 240
 QY 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTTPMSEYEKEVEEMKHMTRQYVASLR 300
 Db 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTTPMSEYEKEVEEMKHMTRQYVASLR 300
 QY 301 RKSSGFSRGASIRYGVTRHHQGRWQARIGRVAGNKDLYLGTFTGQEEAAAYDIAAIKF 360
 Db 301 RKSSGFSRGASIRYGVTRHHQGRWQARIGRVAGNKDLYLGTFTGQEEAAAYDIAAIKF 360
 QY 361 RGLTAVTNFDMNRNVKAILSPISLPISGAARLKEANRPVPSMMTISNNVSESENSASG 420
 Db 361 RGLTAVTNFDMNRNVKAILSPISLPISGAARLKEANRPVPSMMTISNNVSESENSASG 420
 QY 421 WNAVAQHOGVDLSLLHQHQRNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480
 Db 421 WNAVAQHOGVDLSLLHQHQRNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480
 QY 481 TNIDHQSSVSDSVTCGVNMGYGGYQGFAPVNCDAVAASEFDYNARNHYFFAQOQQTQ 540
 Db 481 TNIDHQSSVSDSVTCGVNMGYGGYQGFAPVNCDAVAASEFDYNARNHYFFAQOQQTQ 540
 QY 541 QSPGGDFPAAMTNNVGSNNMYHGGGGVAPFTTWNNDN 579
 Db 541 HSPGGDFPAAMTNNVGSNNMYHGGGGVAPFTTWNNDN 579

RESULT 3

Q8LSN1 PRELIMINARY; PRT; 584 AA.
 ID Q8LSN1
 AC Q8LSN1
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE BABY BOOM.
 GN BBM.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C24;
 RA Boutiller K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,
 RA Ouellet T., Zhang L., Hattori J., Liu C.-M., Miki B.L.A.,
 RA Custers J.B.M., van Lookeren Campagne M.M.;
 RT "Ectopic expression of the Brassica napus BABY BOOM gene triggers a
 RT conversion from vegetative to embryonic growth.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317907; AAM33803.1;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.

NCBI_TaxID=4530;
 [1]
 QY SEQUENCE FROM N.A.
 RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
 RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
 RA Shao Y., Sun Y., Hu Q.F., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
 RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
 RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J.,
 RA Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
 RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
 RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
 RA Zhang R.Q., Guan J.P., Hong G.F.,
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 QY SEQUENCE FROM N.A.
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Wang Q.J., Zhang L.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.,
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL606604; CAB02944.1; --
 DR EMBL; AL731638; CAB0555.1; --
 SQ SEQUENCE 555 AA; 67291 MW; D6FCCD0F729A6933 CRC64;
 Query Match 36.5%; Score 1139; DB 10; Length 655;
 Best Local Similarity 41.6%; Pred. No. 4.4e-69;
 Matches 273; Conservative 69; Mismatches 144; Indels 170; Gaps 21;
 QY 3 NNWLGFSLSPY-EQNHRRKDVSTTTTVDVAGEYCYDPTAASDESSAQTSPSPFV 61
 DB 5 DNWLGFSLSGQNPQHONGQSPSAAGDAADISGDFYGLPTDAHHGNAGEDAYG- 63
 QY 62 VDAFTDRNNSHRDWDINGCACNNIHNDQEQPKLENFLGRITTIYNTNENVGDG- 102
 DB 64 VMDAENRGTHT-ETQDWARNGLDYGGSSDLMLVSSGCGGRRTVAGDVGAPKLENFLD 122
 QY 103 RTTITYNTNENVG-----DGS-----GGCYGGDGGGSLGSLMKTWLRN-----OPVD 146
 DB 123 GNSFSDVHCQAAGGYLSCSAVGGAGGYNGCGGGTIELSMKTWLRNSQOQPPQ 182
 QY 149 NVDCNENGNA-----KGLSLMNSST----- 170
 DB 183 HADQCMSTDAGASSVACSDVLVSCGGGGAGGTASSHCGGLALSMSTGSAAGGGGAVV 242
 QY 171 SCNNNDNNNVAAQGTIDPSVEATPKTIESGQRTSIYRGVTRHRWTRGYEAHLWDN 230
 DB 243 AAESSSSSENKRVDSPGGAVDGA---PRKSIDTGTQRTSIYRGVTRHRWTRGYEAHLWDN 299
 QY 231 SKREGQTEKRGQVYLGQYDKEEKAARAYDLAALKYGTITTTNFPMSYEKEVEEMKHM 290
 DB 300 SCRREGQKRGKQ-----GGYDKEDKAARAYDLAALKYGTITTTNFPMSYEKEVEEMKHM 356
 QY 291 TRQEVASLRKSSGFSRGASIRGVTRHHQGRWQARIGRVAGNKDYLGTFTGTQEEAA 350
 DB 357 TRQEVIAHLRNSGFSRGASKRGVTRHHQGRWQARIGRVAGNKDYLGTFTGTQEEAA 416
 QY 351 EAYDIAAKFRGLTRAVTFDMNRNVKALIESPSLPTGSAARLKEA----- 397
 DB 417 EAYDIAAKFRGLTRAVTFDMNRVDVKSILDSSTLPVGGAAARLKEAEVAAAAAGGVIV 476
 QY 398 -----NRVPVPMNMSINNVSESENASGCV-----QNAAV-Q 427
 DB 477 SHLADGGVGGYGGCGPTIATFGGGGQAPLAVHY-----PSYQASGCKPEQDAVIAA 531
 QY 428 HHQGVDSLILH-----QHORYNGYIYNGNLSSSARACFQEDDQHHFLS 474
 DB 532 GHCATDLQHLHLSGGAAATNFQQPASSAVYCGNGGGGCGNA-----FMW 578
 QY 475 NTQSLMTNIDH---QSSV---SDDSVTVCGNVVVGCGYGGFAAPVNCNDYAAASEFTY 525
 DB 579 PNGAVVAADHGQSSAGGSGDESGRL---VVGVDGV-----VDPYAAMRSAY 623
 RESULT 6
 Q8S1E1 PRELIMINARY; PRT; 692 AA.
 AC Q8S1E1;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative ovule development protein aintegumenta-like protein.
 GN P0035F12.3.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0035F12.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003313; BAB89946.1; --
 DR Gramene; Q8S1E1; --
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 2.
 DR PRINTS; PR00367; ETHRSPLEMT.
 DR ProDom; PD001423; TF_ERF; 2.
 DR SMART; SM00380; AP2; 2.
 SQ SEQUENCE 692 AA; 71515 MW; 4D5A0B49ED8772AF CRC64;
 Query Match 36.3%; Score 1131; DB 10; Length 692;
 Best Local Similarity 40.1%; Pred. No. 1.7e-68;
 Matches 278; Conservative 65; Mismatches 125; Indels 226; Gaps 23;
 QY 3 NNWLGFSLSPYEQNHRRKDVSTTTTVDVAGEYCYDPTAASDESSAQTSPSPFV 62
 DB 5 NNWLAFLSPQQLPP-----SQTNSTLTISAAA-----TTTTAGSDSTGCVCFNIP- 50
 QY 63 VDAFTDRNNSHRDWDINGCACNNIHNDQEQPKLENFLGRITTIYNTNENVGDG- 117
 DB 51 -----QDWSMRGSELISALVAE-----PKLEDFLGISFSEQHHHGKGVIPS 94
 QY 118 SGSGCY----- 128
 DB 95 SAAACVASSGSSGVLYLPPPPSSSLQFADSVNVATSSPVVAHDGVSGGVMSAAAAAAS 154
 QY 129 GCGSLGLSMIKTWLRNQPVDVNDQENGAAGKLSLMSNS----- 168
 DB 155 GNGGIGLSMKWLRNSQPAP-----QPAQLSLMNMAGTTTAQGGGAWALLAGAGE 206
 QY 169 -----STSC-----DNNNDNNNVAAQGTIDDS-----V 193
 DB 207 RGETTPASSLSLSTSAHGATTATWAGGRKEINEEGSGSAGAVVAVGSEGGGAVVEAGAA 266
 QY 194 EATPKTIESFGQRTSIYRGVTRHRWTRGYEAHLWDNSCKREGQTRKRGQVYLGQYDKEE 253
 DB 267 AAAARKSVDTFGQRTSIYRGVTRHRWTRGYEAHLWDNSCRREGQTRKRGQ---GGYDKEE 323
 QY 254 KAARAYDLAALKYGTITTTNFPMSYEKEVEEMKHMTRQEVASLRKSSGFSRGASII 313
 DB 324 KAARAYDLAALKYGTITTTNFPVNNYEKLEEMKHMTRQEFVASLRKSSGFSRGASII 383
 QY 314 RGVTRHHQGRWQARIGRVAGNKDYLGTFTGTQEEAAAYDIAAKFRGLTRAVTFDMNR 373
 DB 384 RGVTRHHQGRWQARIGRVAGNKDYLGTFTGTQEEAAAYDIAAKFRGLTRAVTFDMNR 443

```

QY 374 YNVAKLESPLPIGSAKRLKEANRVP-SMMMISNNVS-----ESENSASGW 421
D 444 YDVKSLDSALPVGTAAKSLKDAEAAAAYDVGRITASHLGGDCAYAAHYGHHHSAAAA 503
QY 422 QNAAVQ-----HHQGVDSLHGHQERVNGYNGNLSSESACFKQEDDQHFLS 474
D 504 PTAFAQAAAAPPAAAG-----LYHPTAOPLRG-----C-KQOD-HAVIA 543
QY 475 NTQSLMTNIDHQSVDSDSVTCGNVGVGYGGAAPVNCDAAYASEFDYNARNHYFA 534
D 544 AAHSQ-DLHH-----LNLGAAAA-----HDFS 567

QY 535 QOQOQTQSPGDPFPAAMTNVGSN-MYHGEGBG 567
D 568 QAWQOQHGLGSDNASLEHSTGNSVYVNGDNG 601

RESULT 7
Q8LQ3 PRELIMINARY; PRT; 597 AA.
ID Q8LQ3
AC Q8LQ3;
DT 01-OCT-2002 (TremBrel. 22, Created)
DT 01-OCT-2002 (TremBrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBrel. 24, Last annotation update)
DE Ovule development aintegumenta-like protein BNM3.
GN BNM3.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
CX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Bl X-Z;
RT "Cloning and identification of two ovule development proteins,
RT aintegumenta-like protein in rice (Oryza sativa).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY062180; AAL47205.1; -
DR Gramene; Q8LQ3; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR PRINTS; PR00367; ETHRSPLEMT.
DR ProDom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2_2.
SQ SEQUENCE 597 AA; 62198 MW; P856BEC99BADE25B CRC64;

Query Match 34.9%; Score 1089.5; DB 10; Length 597;
Best Local Similarity 41.2%; Pred. No. 9e-66;
Matches 264; Conservative 54; Mismatches 109; Indels 213; Gaps 20;

QY 3 NNWLGSLSPYEQNHHRKDVYSSTTTVDVAGEYCDPTAASDESSAIQTSFSPFGVV 62
D 5 NNWLAFLSPQDQLPP-----SQNSTFISAAA-----TTTAGDSST----- 42
QY 63 VDAFTRDNNSHRWDINGCACNNIHNDQDGPKLENFLGRTTIYNTNENVDGSGSGC 122
D 43 -----GDVCFNIPOAHPSTPAI----- 59
QY 123 YGGDGGGSLGLSMIKTFLRNQPVNDVNGENGAAGLSLMSNS----- 168
D 60 -----GNGGIGLSMINKLRSQAP-----QPAQALSLSMNWAGTTTAQGGGAMAL 105
QY 169 -----STSC-----DNNNDNNNVAAQKTDIDS--- 192
D 106 LAGAGERGRTTPASESLSTSAGATTATMAGRKEINEEGSGAGAVVVGSSGSGAV 165
QY 193 -----VEATPKKTTIESGQRTSIYRGVTRHRWTCRYEAHLWDSCKREGQTRKGRQYILG 247
D 166 VEAGAAAAAARKSVDTTGQRTSIYRGVTRHRWTCRYEAHLWDSCKREGQTRKGRQ---G 222

```

```

QY 248 GYDKEEKAARAYDLAALKYWGTTTTTNPMPSEYKEVEEMKMTQEVVASLRKSSGFS 307
D 223 GYDKEEKAARAYDLAALKYWGTTTTTNPMPSEYKEVEEMKMTQEVVASLRKSSGFS 282
QY 308 RGASIRGVTRHHQGRWOARIGRVAGNKDLYLGTFTGQBEAABAYDIAAIKFRGLTAVT 367
D 283 RGASIRGVTRHHQGRWOARIGRVAGNKDLYLGTFTGQBEAABAYDIAAIKFRGLNAV 342
QY 368 NFDMMRYNKAILESPLPIGSAKRLKEANRVP-SMMMISNNVS-----ESE 415
D 343 NFDMSRYDVKSLDSALPVGTAAKSLKDAEAAAAYDVGRITASHLGGDCAYAAHYGHHH 402
QY 416 NSASQWQNAVQ-----HHQGVDSLHGHQERVNGYNGNLSSESACFKQEDD 468
D 403 SAAAAPTATACAAAAPPAAAG-----LYHPTAOPLRG-----C-KQOD 443
QY 469 QHFLSNTQSLMTNIDHQSVDSDSVTCGNVGVGYGGAAPVNCDAAYASEFDYNAR 528
D 444 HAVIAAAHSQ-DLHH-----LNLGAAAA----- 467

QY 529 NHYPAQOQOQTQSPGDPFPAAMTNVGSN-MYHGEGBG 567
D 468 -HDFSQAMQOQHGLGSDNASLEHSTGNSVYVNGDNG 506

RESULT 8
Q9SYC2 PRELIMINARY; PRT; 516 AA.
ID Q9SYC2
AC Q9SYC2;
DT 01-MAY-2000 (TremBrel. 13, Created)
DT 01-MAY-2000 (TremBrel. 13, Last sequence update)
DT 01-JUN-2003 (TremBrel. 24, Last annotation update)
DE FlM15.6 protein.
GN FlM15.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A.; Palm C.J.; Conway A.B.; Conn L.; Hansen N.F.;
RA Altafi H.; Araujo R.; Huizar L.; Rowley D.; Buehler E.; Dunn P.;
RA Gonzalez A.; Kremenetskaia I.; Kim C.; Lenz C.; Li J.; Liu S.;
RA Luros S.; Schwartz J.; Shinn P.; Toriumi M.; Vysotskaia V.S.;
RA Walker M.; Yu G.; Ecker J.; Theologis A.; Davis R.W.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005085; AAD30633.1; -
DR FIR; F96549; F96549.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR PRINTS; PR00367; ETHRSPLEMT.
DR ProDom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2_2.
SQ SEQUENCE 516 AA; 56335 MW; 80996A95851B8BEA CRC64;

Query Match 33.4%; Score 1040.5; DB 10; Length 516;
Best Local Similarity 44.6%; Pred. No. 1.6e-62;
Matches 254; Conservative 70; Mismatches 114; Indels 131; Gaps 22;

QY 86 NIHNDQDQ---PKLENFLGRTTT-IYNTNEN-----VGDGSGGCGYGGDGGGSLGL 135
D 3 NPHGGGEGGEVPAKADFLGVSKSGDHHHTDHLVFNDIHQTNASDYVFQTN-----SL-L 57
QY 136 SMIKTFLRNQPVNDVNGENGAAGLSLMSNS-----TSCDNNNDN 179
D 58 PTVVCASNAP-NNVELOESAHNLQSLTSLMSGTGAAAEVATVKAAPETSADSSSTT 116
QY 180 NNWVAQKTDIDSVEATPKKTTIESGQRTSIYRGVTRHRWTCRYEAHLWDSCKREGQTR 239

```



```
DR SMART: SMO0380; AP2; 2.
SQ SEQUENCE 566 AA; 63257 MW; 470627FB3002744A CRC64;

Query Match 32.0%; Score 997; DB 10; Length 566;
Best Local Similarity 40.4%; Pred. No. 1.6e-59;
Matches 247; Conservative 58; Mismatches 127; Indels 180; Gaps 18;

QY 43 AASDESSAIQTSFSPFGVVDAETRDNNSHSRDWDINGCACNNIHNDQDGPGLNPLG 102
DB 52 AASVADSTLITFVDP-----OSHSQNHIPKLEDFLG 84

QY 103 RT-----TTIY-----NTNEN 113
DB 85 DSSGIVYSDNSQTDQDSSLTQIYDPRHHNQGFYSDHDDFKTMAGFQSAPSTNSGS 144

QY 114 VDGSGSG-----CYGGGCGGSGSLSMIKTWLRNQPVDVNDQ 153
DB 145 VDDSGASGTHLAGDYLGHVYESSGPGLGFHGGSGALSLGVNNNN--TNHRDN-DNH 201

QY 154 ENGNAAGLSLMSNSTSCDNNNDNNNVVAQK---TIDDSVEATPKKTIESFGQRTS 209
DB 202 YRGN-----NNGERINNNNDNEKTDSEKEKAVVAVEISDCSNKKIADTFGQRTS 252

QY 210 IYRGVTRHRTGRYEHLWDNSCKREGQTRKGRQVYLGVDKBEKAARAYDLAALKYWG 269
DB 253 IYRGVTRHRTGRYEHLWDNSCRREGQARKGRQ---GGYDKEDKAARAYDLAALKYWN 309

QY 270 TTTNFPMSYEKEVEEMKMTROEYVASLRKSSGFSRGASIVRGVTRHHQGRWOARI 329
DB 310 TATTNFIITNYSKEVEEMKMTKEBFIASLRKSSGFSRGASIVRGVTRHHQGRWOARI 369

QY 330 GRVAGNKDLVLTGFTQBEAAEAYDIAAKFRGLTAVTNFDMNRYNKAILESPLSIG 389
DB 370 GRVAGNKDLVLTGFTATBEAAEAYDIAAKFRGINAVTNFDMRYDVEALMK-S-ALPI 428

QY 390 AAKELKEANRPVPSMMNISNNVSESENASGQWNAAVQHQQGVDSLHLLHQERYNGY 449
DB 429 AAKELK-----LSLEAAASSEOKFILGHQ-----LHFQOQO----- 461

QY 450 NGGNLSSESARACPKQEDDQHHFLSNLQSLMTNIDHQSVDSDSVTCVGNVVGYYG 509
DB 462 -----QQOQLQCSSPNHSSINFALCPNSAVSQIIFCG----- 496

QY 510 AAPVNCDAAYASEFDYNARHYHFAQQCQTSQSPGGFPA-AMTNVGNVNNYHVBEGG 567
DB 497 ---TFEAAALYHHHQOQOCH---QQOQOQOQFFQHFPAANAASDSTGSNNNSNV 549

QY 568 EVAPT---FTW 576
DB 550 LMAENPAEFFLM 561

RESULT 11
Q84202 ID Q84202 PRELIMINARY; PRT; 639 AA.
AC Q84202;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 25, Last annotation update)
DE Putative ovule development protein antitegumenta (ANT).
GN P0474G09.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0474G09.16";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AP005309; BAC56815.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001471; TF_ERF.
DR Pfam: PF00847; AP2-domain; 2.
DR PRINTS: PR00367; ETRHSPELEWMT.
DR PRODOM: PD001423; TF_ERF; 2.
DR SMART: SMO0380; AP2; 2.
SQ SEQUENCE 639 AA; 67632 MW; EE4BC5123C56E68F CRC64;

Query Match 31.1%; Score 971.5; DB 10; Length 639;
Best Local Similarity 36.3%; Pred. No. 1e-57;
Matches 252; Conservative 79; Mismatches 174; Indels 189; Gaps 22;

QY 3 NNWLGSLSPY-----BQNHHRKD-----VYSSTTTTVDVAVAG 35
DB 16 SGNLGSLSFHHMASSTMDSHHHVHHHQQOQQOQQOQQHQQOQGLFPFVSTTAAAAAY 75

QY 36 EYCYDPTASDESSAIQTSPSPFG--VVVDAFTRDNNSHSRDWDINGCACNNIHNDQ 92
DB 76 GLAGDVVAATNGYYSQLASMFLKSGSLCIMEALRRTDQDH----- 116

QY 93 DQPKLENPLG-----RTTIYNTNENVDGSGGCGYGGDGG-----GSLGLSMI- 138
DB 117 HGFPLEDFLGAAQAPAMALSLDNTSGFYGGGCAAGAGHGHQGYLQACDLYGGPAPSLVT 176

QY 139 -----KWL-ENQVDNVNDQENGNAAK-----GLSLSMN 167
DB 177 ADEEAAAAAANAASVAAARGAATAYATGAADANAENVLPSATAAQHLHFLALSMSSG 236

QY 168 SSTSDNNNDNNNVVA-----QGKTIDDSVEATPKKTIESFGQRTSIYR 212
DB 237 SLSSCITACEYGMVAADADGGRKGGAGGQKQPVHH-----RKSIDTFQRTSQYR 290

QY 213 GVTTRHRTGRYEHLWDNSCKREGQTRKGRQVYLGVDKBEKAARAYDLAALKYWGTTT 272
DB 291 GVTTRHRTGRYEHLWDNSCKREGQTRKGRQ---GGYDMEEKAAARAYDLAALKYGPSTH 347

QY 273 TNFPMSYEKEVEEMKMTROEYVASLRKSSGFSRGASIVRGVTRHHQGRWOARI 332
DB 348 INFPLEDQEELEEMKMTROEYVAHLARKSSGFSRGASIVRGVTRHHQGRWOARI 407

QY 333 AGNKDLYLTGFTQBEAAEAYDIAAKFRGLTAVTNFDMNRYNKAILESPLSIGSAK 392
DB 408 SGNKDLVLTGFTQBEAAEAYDIAAKFRGLNAVTFDITRYDVKIWNASNTLLPADLAR 467

QY 393 RLKEANRPVPSMMNISNNVSESENASGQWNAAVQHQQGVDSL-----HQQE 442
DB 468 R-----NAATTSKDDHSAAG-AGAIVSVHSAADIADVADTLWKATAPRQOQO 514

QY 443 RYNGYINGGN-----LSSESARACPKQEDDQHHFLSNLQSLMTNIDHQSVDSD 492
DB 515 HHDDVVLGADQAFSVLHDLVAVDAAAHHQOQQOQQHMSMSAASSLVTSLNSREGSPD 574

QY 493 SVTVCGNVVGYGQGFAPVNCDAAYASEFDYNARHYHFAQQCQTOO-----SPG 547
DB 575 R-----GGLSMLFAKP--SPVAAS-----AQOQASTKLMAAPLPLGWS 612

QY 548 ---PAAMTNVGNVNNYHVBEGGGEVAPTFTWMD 578
DB 613 VSSPPASAPPPAVSIAH-----MPLFAAWTD 638

RESULT 12
Q8LQ04 ID Q8LQ04 PRELIMINARY; PRT; 492 AA.
AC Q8LQ04;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aintegumenta-like protein.
OS Oryza sativa (Rice).
```



```
RESULT 14
Q41832
ID Q41832 PRELIMINARY; PRT; 485 AA.
AC Q41832;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Orf protein.
GN ORF.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Black Mexican Sweet; TISSUE=Endosperm;
RA Daniel T.J. Fordham-Skelton A.P., Vergani P., Edwards R.;
RT "Isolation of a Maize cDNA (Accession No. Z47554) Encoding APTALA2-
RT Like Binding Domains by Complementation Cloning of an L-Isoaspartyl
RT Methyltransferase-Deficient Mutant of Escherichia coli (PGR 96-013).";
RL Plant Physiol. 110:1435-1435(1996).
DR EMBL; Z47554; CAA87634.1; -.
DR PIR; T03638; T03638.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR ProDom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2_2.
DR SEQUENCE 485 AA; 52391 MW; 2B165E3D17FC8B3A CRC64;
QY
Query Match 29.8%; Score 930; DB 10; Length 485;
Best Local Similarity 40.0%; Pred. No. 4.8e-55;
Matches 233; Conservative 51; Mismatches 131; Indels 168; Gaps 19;
Db
3 NNMGLPSLSPYEQNHHR--KDVYSTTTTVDVAGEYCYDPTAASDESSAIQTSFSPFG 60
11 HHWLSFSL--NNYHGLEAFNSSGTFL-----GDEQGAVER-----48
61 VVVDFTRDNNSHSRDWDINGCACNNIHNDQDQPK-LNPLGRITTYNTNENVGSG 119
49 -----PRTVDFLG-----G 58
120 SGCVGGDGG-----GSLGLSNKWTNRQVNDVNDQENGNAKGLSLSNSS 169
59 VGCVGAPRSRLQIRITSLCAASCG-SITARFLRHYPAAQ-SGTTVGEPLSRFTLAAMS 116
170 TSCDNNDSNNVVAQKTTIDDSVEATPKKTIESGQRTSIYRGVTRHRTGRYEAHLWD 229
117 TD-----VAAESDQAS-----RSAETFGQRTSIYRGVTRHRTGRYEAHLWE 159
230 NSCRREGTRKGRQVYLGVDYKSKAARAYDLAALKTWGTTTTFPMSEYEKEVEBMKH 289
160 NSCRREGTRKGRQVYLGVDYKSKAARAYDLAALKTWGTTTTFPMSEYEKEVEBMKS 219
290 MTRQEVASLRKSGSPSRGASIRGVTRHHQHWQARIGRVAGNKDLYLGFSTQBEA 349
220 MTRQEVASLRKSGSPSRGASIRGVTRHHQHWQARIGRVAGNKDLYLGFSTQBEA 279
350 AEAYDIAAKPRGLTAVTNFDMRYNKAILESPLFGSAKRLKEANRPVPSM-----404
280 AEAYDIAAKPRGLTAVTNFDMRYNKAILESPLFGSAKRLKEANRPVPSM-----338
405 --NMISNNVSE-----NSASGQNAVOHOGVDLSLLHCHOER-----YNGYYNG 451
339 AAMMLAGAAASQATMPPEKDYKWSLLAHYQQ-----QQEQRFPPASAYEAYSGG 391
452 GNLSSESARACFKQEDDQHHFLSNTQSLMTN-----IDHQSSVSDSDSVTV 496
392 VNVD-----FTMGTSSGNNNTGSGVMWGATSGAVVQQDDSSSKQNGY 435
```

```
QY 497 CGN-----VVGYGVGQGFAPVNCDAAYAASEFDYNARNHY 532
Db 436 ASNPYAAAMVSGTAGYEGSTGDNGTWTTTTSNTGTAPHY 478

RESULT 15
Q9CAH3
ID Q9CAH3 PRELIMINARY; PRT; 425 AA.
AC Q9CAH3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Putative AP2 domain transcription factor.
GN F28F22.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.B., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltcher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzig S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC010926; AAG51860.1; -.
DR PIR; B96750; B96750.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR PRINTS; PR00367; ETRSPLEMMNT.
DR ProDom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2_2.
DR SEQUENCE 425 AA; 48205 MW; C191BA52D65ACBB CRC64;
QY
Query Match 29.7%; Score 926; DB 10; Length 425;
Best Local Similarity 45.0%; Pred. No. 7.5e-55;
Matches 207; Conservative 47; Mismatches 86; Indels 120; Gaps 12;
Db
5 WLGFSLSP-----YEONHRRKDVYSTTTTVDVAGEYCYDPTAASDESSAIQTSFSP 58
4 WLGFSLTPPLRICNSEEBELRHGSDGTL-----32
59 FGVVDAFTRDNNSHSRDW--DINGCACNNIHNDQDQPKLENFLGR-----104
33 -----FFTH-----FDPVVRWDIN-----FDPHHHDEDEKVEDLLSNSTYPTINMHT 78
105 -----TTTYNTNENVG-----DGSQSGCVGGDGGGSLGLSMKLTWLR- 143
79 NVNCTTVNRLNPPGVLHDDQVTPHYPLNPLNDYGGFERVG---SVSFKSLWLEQ 135
144 -----NQPVDNVDNQENGNAAG--LSSLNMSSTSCDNNDSNNV- 182
136 GTFAPFLSHSYVTEAGTGNISHSFNEETGYNTGMSMLSLSHGACDLSNESVSAR 195
```

```
QY 183 VAGKTIIDS-----VEATPKTIESFGORTSYRGVTRHRTGRYEAHLWDNSC 232
Db 196 VEEPVKDEKRLVVKPOKESVPRKSVDSYGQTSQYRGVTRHRTGRYEAHLWDNSC 255
QY 233 KREGQTRKGRQVYLGGYDKKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKHMTR 292
Db 256 KKEGQTRGRQAKIGGYDEEKAARAYDLAALKYWGPTTHLNFPLSNYEKEIEELNNMR 315
QY 293 QEYVASLRKSSGFSRGASIVGVTTRHHQHGHGWQARIQGVAGNKDLYLGTGTQEEAEEA 352
Db 316 QEFVAMLRNNSGFSRGASVYRGVTRHHQHGHGWQARIQGVAGNKDLYLGTGTQEEAEEA 375
QY 353 YDIAAIKFRGLTAVTFNFMNRYNVKAILSPSLPICSAK 392
Db 376 YDIAAIKFRGLNAVTFNFDINRYDVVKRICSSSTIVDSQAK 415
```

Search completed: July 15, 2004, 10:03:10
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 09:51:59 ; Search time 57 seconds
(without alignments)
2870.088 Million cell updates/sec

Title: us-09-980-364-2

Perfect score: 3119

Sequence: 1 MNNNLGFLSLSPYEQNHRRK.....YYHGGGGEVAPTFTWINDN 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3119	100.0	579	4 AAB50873	Aab50873 Brassica
2	3119	100.0	579	4 AAB50872	Aab50872 Brassica
3	1310	42.0	707	6 ABR40766	Abr40766 Glycine m
4	1296	41.6	710	6 ABR40768	Abr40768 Glycine m
5	1131	36.3	692	6 ABR40877	Abr40877 Oryza sat
6	1119.5	35.9	553	6 ABR40771	Abr40771 Glycine m
7	1075.5	34.5	486	6 ABR40767	Abr40767 Glycine m
8	1047.5	33.6	574	7 ADD30879	Add30879 Plant yie
9	1047.5	33.6	574	7 ADD30334	Add30334 Plant yie
10	1047.5	33.6	574	7 ADD30823	Add30823 Plant yie
11	1040.5	33.4	516	6 ABR40836	Abr40836 Arabidops
12	1011	32.4	585	5 ABB79640	Abb79640 Cotton Al
13	1010.5	32.4	538	6 ABR40854	Abr40854 Gossypium
14	1003	32.2	470	6 ABR40853	Abr40853 Catalpa s
15	999	32.0	512	6 ABR40852	Abr40852 Glycine m
16	997.5	32.0	540	6 ABR40837	Abr40837 Arabidops
17	997	32.0	566	6 ABR40876	Abr40876 Arabidops
18	994	31.9	557	6 ABR40856	Abr40856 Glycine m
19	986	31.6	510	6 ABR40764	Abr40764 Glycine m
20	977	31.3	558	7 ADD30981	Add30981 Plant yie
21	974.5	31.2	498	7 ADD30350	Add30350 Plant yie
22	971	31.1	530	6 ABR40769	Abr40769 Glycine m
23	964.5	30.9	469	6 ABR40759	Abr40759 Oryza sat
24	961	30.8	492	6 ABR40855	Abr40855 Zea mays
25	957.5	30.7	384	6 ABR40770	Abr40770 Glycine m

26	951.5	30.5	642	5 ABB79638	Abb79638 Rice AINT
27	949.5	30.4	484	6 ABR40757	Abr40757 Zea mays
28	943.5	30.3	489	6 ABR40756	Abr40756 Zea mays
29	942	30.2	663	5 ABB79636	Abb79636 Soybean A
30	942	30.2	669	5 ABB79639	Abb79639 Rice AINT
31	941	30.2	415	7 ADD30560	Add30560 Plant yie
32	930	29.8	485	6 ABR40829	Abr40829 Zea mays
33	930	29.8	485	6 ABR40834	Abr40834 Zea mays
34	926	29.7	425	6 ABR40833	Abr40833 Arabidops
35	924	29.6	665	5 ABB79637	Abb79637 Soybean A
36	913.5	29.3	548	3 AAB07725	Aab07725 A Brassic
37	891	28.6	555	3 AAB07724	Aab07724 An Arabid
38	891	28.6	555	3 AAG31380	Aag31380 Arabidops
39	891	28.6	555	6 ABR40830	Abr40830 Arabidops
40	889	28.5	555	6 ABR40832	Abr40832 Arabidops
41	880	28.2	528	3 AAG31382	Aag31382 Arabidops
42	880	28.2	529	3 AAG31381	Aag31381 Arabidops
43	714	22.9	312	6 ABR40762	Abr40762 Glycine m
44	711.5	22.8	409	6 ABR40849	Abr40849 Momordica
45	681	21.8	345	6 ABR40875	Abr40875 Arabidops

ALIGNMENTS

RESULT 1
AAB50873
ID AAB50873 standard; protein; 579 AA.

AC AAB50873;

DT 19-MAR-2001 (first entry)

DE Brassica napus BMN3B protein.

XX Brassica napus microspore embryo; BMN; microspore embryogenesis; BMN3B;
XX asexually derived embryo production; increased regenerative capacity.

OS Brassica napus.

PN EP1057891-Al.

PD 06-DEC-2000.

PF 02-JUN-1999; 99EP-00201745.

PR 02-JUN-1999; 99EP-00201745.

XX (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
XX (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
PI Boutillier K;

XX WPI; 2001-026720/04.
XX N-PSDB; AAC91399.

PT Brassica napus BMN3 DNA sequences, useful for transforming host cells to
PT produce asexually derived embryos.

PS Claim 62; Fig 3; 51pp; English.

XX The present sequence is given in a specification relating to a Brassica
XX napus microspore embryo (BMN) gene obtained during the induction of
XX microspore embryogenesis. The isolated DNA molecule can be used in a
XX vector that is used to transform plant cells. The vector can be used in
XX methods to produce asexually derived embryos, for modifying the
XX regenerative capacity of a plant, and for producing an apomictic plant
XX Sequence 579 AA;

Query Match 100.0%; Score 3119; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 3e-257;

Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNWLGFSLSPEQNHHRKDVYSSTTTTVDVAGEYCYDPTAADESSAIQTSPFPFG 60
 Db 1 MNNWLGFSLSPEQNHHRKDVYSSTTTTVDVAGEYCYDPTAADESSAIQTSPFPFG 60

QY 61 VVDAFTDRNNSHRDWDINGCACNNHNDQDGPKEFLGRITTIYNTNENVDGSGS 120
 Db 61 VVDAFTDRNNSHRDWDINGCACNNHNDQDGPKEFLGRITTIYNTNENVDGSGS 120

QY 121 GCGGGGGGSLGLSMIKTWLRNPVDVNOENGAAGLSLMSNSTSCDNNNSNN 180
 Db 121 GCGGGGGGSLGLSMIKTWLRNPVDVNOENGAAGLSLMSNSTSCDNNNSNN 180

QY 181 NVVAQKTIIDDSVEATPKKTIESFQRTSIYRGVTRHWTGRYEHLWDNSCKREGQTRK 240
 Db 181 NVVAQKTIIDDSVEATPKKTIESFQRTSIYRGVTRHWTGRYEHLWDNSCKREGQTRK 240

QY 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTNPMSYEKEVEEMKHMTRQEVVASLR 300
 Db 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTNPMSYEKEVEEMKHMTRQEVVASLR 300

QY 301 RKSSGFSRGASIRYGVTRHWHQGRWQARI GRVAGNKDLYLGTFTGQEEAAAYDIAAKF 360
 Db 301 RKSSGFSRGASIRYGVTRHWHQGRWQARI GRVAGNKDLYLGTFTGQEEAAAYDIAAKF 360

QY 361 RGLTAVTNFDMNRNVKAILSPSPIGSAARLKEANRPVPSMMISNNVSESENSASG 420
 Db 361 RGLTAVTNFDMNRNVKAILSPSPIGSAARLKEANRPVPSMMISNNVSESENSASG 420

QY 421 WQNAVQHHQGVLDLSLLHQHQRVNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480
 Db 421 WQNAVQHHQGVLDLSLLHQHQRVNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480

QY 481 TNIDHQSSVSDSVTCGVNGYGYQGFAAPVNCDAYAASEFDYNARNHYFFAQOQQTQ 540
 Db 481 TNIDHQSSVSDSVTCGVNGYGYQGFAAPVNCDAYAASEFDYNARNHYFFAQOQQTQ 540

QY 541 QSPGGDFPAAATNNVGSNNMYHGGEGGEVAPFTTVMNDN 579
 Db 541 QSPGGDFPAAATNNVGSNNMYHGGEGGEVAPFTTVMNDN 579

RESULT 2
 ID AAB50872 standard; protein; 579 AA.
 AC AAB50872;
 DT 19-MAR-2001 (first entry)
 DE Brassica napus BNM3A protein.
 KW Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3A;
 KW asexually derived embryo production; increased regenerative capacity.
 OS Brassica napus.
 PN EP1057891-A1.
 PD 06-DEC-2000.
 PF 02-JUN-1999; 9SEP-00201745.
 PR 02-JUN-1999; 9SEP-00201745.
 PA (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
 PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
 PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
 PI Boutillier K;
 XX WPI; 2001-026720/04.

DR N-PSDB; AAC91398.
 XX Brassica napus BNM3 DNA sequences, useful for transforming host cells to
 PT produce asexually derived embryos.
 XX Claim 61; Fig 3; slpp; English.
 CC The present sequence is given in a specification relating to a Brassica
 CC napus microspore embryo (BNM) gene obtained during the induction of
 CC microspore embryogenesis. The isolated DNA molecule can be used in a
 CC vector that is used to transform plant cells. The vector can be used in
 CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant
 XX Sequence 579 AA;

Query Match 100.0%; Score 3119; DB 4; Length 579;
 Best Local Similarity 100.0%; Pred. No. 3e-257;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNWLGFSLSPEQNHHRKDVYSSTTTTVDVAGEYCYDPTAADESSAIQTSPFPFG 60
 Db 1 MNNWLGFSLSPEQNHHRKDVYSSTTTTVDVAGEYCYDPTAADESSAIQTSPFPFG 60

QY 61 VVDAFTDRNNSHRDWDINGCACNNHNDQDGPKEFLGRITTIYNTNENVDGSGS 120
 Db 61 VVDAFTDRNNSHRDWDINGCACNNHNDQDGPKEFLGRITTIYNTNENVDGSGS 120

QY 121 GCGGGGGGSLGLSMIKTWLRNPVDVNOENGAAGLSLMSNSTSCDNNNSNN 180
 Db 121 GCGGGGGGSLGLSMIKTWLRNPVDVNOENGAAGLSLMSNSTSCDNNNSNN 180

QY 181 NVVAQKTIIDDSVEATPKKTIESFQRTSIYRGVTRHWTGRYEHLWDNSCKREGQTRK 240
 Db 181 NVVAQKTIIDDSVEATPKKTIESFQRTSIYRGVTRHWTGRYEHLWDNSCKREGQTRK 240

QY 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTNPMSYEKEVEEMKHMTRQEVVASLR 300
 Db 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTNPMSYEKEVEEMKHMTRQEVVASLR 300

QY 301 RKSSGFSRGASIRYGVTRHWHQGRWQARI GRVAGNKDLYLGTFTGQEEAAAYDIAAKF 360
 Db 301 RKSSGFSRGASIRYGVTRHWHQGRWQARI GRVAGNKDLYLGTFTGQEEAAAYDIAAKF 360

QY 361 RGLTAVTNFDMNRNVKAILSPSPIGSAARLKEANRPVPSMMISNNVSESENSASG 420
 Db 361 RGLTAVTNFDMNRNVKAILSPSPIGSAARLKEANRPVPSMMISNNVSESENSASG 420

QY 421 WQNAVQHHQGVLDLSLLHQHQRVNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480
 Db 421 WQNAVQHHQGVLDLSLLHQHQRVNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480

QY 481 TNIDHQSSVSDSVTCGVNGYGYQGFAAPVNCDAYAASEFDYNARNHYFFAQOQQTQ 540
 Db 481 TNIDHQSSVSDSVTCGVNGYGYQGFAAPVNCDAYAASEFDYNARNHYFFAQOQQTQ 540

QY 541 QSPGGDFPAAATNNVGSNNMYHGGEGGEVAPFTTVMNDN 579
 Db 541 QSPGGDFPAAATNNVGSNNMYHGGEGGEVAPFTTVMNDN 579

RESULT 3
 ABR40766
 ID ABR40766 standard; protein; 707 AA.
 AC ABR40766;
 XX 16-MAY-2003 (first entry)
 DE Glycine max oil trait related protein sequence SEQ ID NO:352.
 DE Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW

KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CAC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant.
 XX Glycine max.
 XX WO2003002751-A2.
 XX 09-JAN-2003.
 XX 27-JUN-2002; 2002WO-US020152.
 XX 29-JUN-2001; 2001US-0301913P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX WPI: 2003-201509/19.
 XX N-PSDB; ACC00801.
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 12; Page 369-371; 542pp; English.
 XX The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, Lip15-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CAC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (II) can also be
 CC used as a hybridisation probe. ACC00826 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX Sequence 707 AA;
 SQ
 Query Match 42.0%; Score 1310; DB 6; Length 707;
 Best Local Similarity 45.6%; Pred. No. 1.4e-102;
 Matches 322; Conservative 77; Mismatches 139; Indels 168; Gaps 29;
 QY 4 NLLGFLSPYEQNHRRKDVYSST-----TTVDVAGEYCVDPAADESSAIQ 52
 DB 5 NLLGFLSPHEPSSQD-HSQITPSPFPNPGSISSTDVAGG-CFDLTS---DSTPHL 59
 QY 53 TSPSPFPGVVDAFTDNN-SHSRDMDIN-----GCACN--NHNDEQDGKLENFL 101
 DB 60 LNLPS-YG-IYEAFHRNNSINTQDWKENYNSQNLGLTSCNQNNQNNQOQQOQKLENFL 117
 QY 102 GRTTITNTNENV-GDGGSGGCY-----GGGDGGG-----GSLGLSMKTLWR 143
 DB 118 GGHG--FGEHQYVGNASTDWFPAQPVSAAGSGGSGGNNNNNSNLSGLSNIKTLWR 175
 QY 144 NOPVDNVNONGNAAG-----LSLSMN-----SST-----SCDNNDSNN-- 181
 DB 176 NQP-PNSNINNNNSGGINRSSVQQTLSLSMSTGSSQSTSLPLLTAASVDNGSSSDNQ 234
 QY 182 ---VVAQGIIDSVEATPKKTIESTQRTSIYRGVTRHWTGCTGYEAHLWDNSCKREGOT 238
 DB 235 PNTSAALDSTQTGAETAPRKSIDTFQRTSIYRGVTRHWTGCTGYEAHLWDNSCKREGOT 294

QY 239 RKGRQVILGGYDKEEKAARAYDLAALKYWGCTTTTTFPMSEYKEVEEMKHTQEVAS 298
 DB 295 RKGRQVILGGYDKEEKAARAYDLAALKYWGCTTTTTFPIHSEYKELEBEMKHTQEVAS 354
 QY 299 LRRKSSGFSRGASIVRGVTRHHQHGCRMQARIGRVAGNKDLVLTGTFGTQEEAAEAYDAAI 358
 DB 355 LRRKSSGFSRGASIVRGVTRHHQHGCRMQARIGRVAGNKDLVLTGTFGTQEEAAEAYDAAI 414
 QY 359 KFRGLTAVTNFDMRYNVKAILSPSLPIGSAARLKE-----ANRPVPSMM 405
 DB 415 KFRGLSAYTNFDMRYDKVKSILESTTLPIGGAARLKDMEQVELSVNDRHQAQVDHSII 474
 QY 406 MISNNVSSSENSASGQNAVAQHGHQGVDSLHLLHQERYNGYYNGGNLSSESARACFKQ 465
 DB 475 MSSHLTQGINNNYAGGCTAT--HHNWHNAHAFHQPQCTTTHYFYGQRIN-----WCKQE 527
 QY 466 EDDQ-----HHFLSNTQSL--MTNIDHQSSVSDSDSVTVCGNVV 501
 DB 528 QQDNSDAPHSLSYSDIHQLQLGNNGTHTNFHTNSGLHFLMSD---SASIDNSSSSNSW 584
 QY 502 --GYGGYOGF-AAPV-NCDAYAASEFDVN-----526
 DB 585 YDYGCGGGYNNMPMGTTTAVVAGDQNPNSNHGFGDNEIKALGYESVYGSATDSYHAF 644
 QY 527 ARNHYYFAQQOQTQOSPGGDFPAAAMNNVGNMYYHGGEGGEVAPT 572
 DB 645 ARNLYYLTQQQ-----SSSVDTVKASAYDQGSACNTWVPT 679

RESULT 4

ABR40768
 ID ABR40768 standard; protein; 710 AA.
 XX ABR40768;
 AC ABR40768;
 DT 16-MAY-2003 (first entry)
 XX Glycine max oil trait related protein sequence SEQ ID NO:356.
 DE Glycine max oil trait related protein sequence SEQ ID NO:356.
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CAC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant.
 XX Glycine max.
 XX WO2003002751-A2.
 XX 09-JAN-2003.
 XX 27-JUN-2002; 2002WO-US020152.
 XX 29-JUN-2001; 2001US-0301913P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX WPI: 2003-201509/19.
 XX N-PSDB; ACC00803.
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 12; Page 374-376; 542pp; English.
 XX The present invention describes an isolated nucleotide fragment (I)
 CC

comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (pp) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, LIP15-like transcription factor activity, calcosin-like activity, ATP citrate lyase activity, SNF1-like activity, and CKC-like transcription factor activity. Also described: (I) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention

XX
SQ Sequence 710 AA;

Query Match 41.6%; Score 1296; DB 6; Length 710;
Best Local Similarity 45.5%; Pred. No. 2.3e-101;
Matches 324; Conservative 66; Mismatches 146; Indels 176; Gaps 29;

QY 4 NWLGSLSPYE-----QNHHR-----KDYGSTTTTVDVAGYCYDPTAASDESSA 50
DB 5 NLLGSLSPQHPSSQDHSQAPRCFNPDISSST-----DVAGD-CFGLTS---DSTP 55
QY 51 IQTSPSPFGVVDAFTDRDNNSH-SRDWDIN-----GCACNN---IHN---DEQDGP 95
DB 56 HLLNLP-S-YG-IYFAFHSRNIHTQDKENYNSQNLGLTSCSNQNNHNNHQOQQQP 113
QY 96 KLENFLGRTTIYNTNEN-VDGGSGGCGY-----GGDGG-----GSLGLSM 137
DB 114 KLENFLGHS--FGSEHQPYGNSASTEYMPAPQVLAGGGGGGSSNSSTNSSSIGLSM 171
QY 138 IKTWLRNPFVNDVNDQNGNAAG-----LSLSNN-----SST-----SCDNNND 177
DB 172 IKTWLRNPFHSENNNNNNNSGNSRSSVQOTLSLSNSTGQSSTSLPLLTASVDNGES 231
QY 178 SNNN-----VVAQGTIDDSVEATPKTIESFGORTSIYRGVTRHRTWGRYEAHLWNSC 232
DB 232 SSDNKQPHTTAALDTQTGA-ETAPRKSIDTFGORTSIYRGVTRHRTWGRYEAHLWNSC 291
QY 233 KREGTRKRGVYLGQYDKKKAARAYDLAALKYWGTTTTTNPMSYEKEVEEMKMTNR 292
DB 292 RREGTRKRGVYLGQYDKKKAARAYDLAALKYWGTTTTTNPFIHYEKELEEMKMTNR 351
QY 293 QEYVASLRKSGSFGSRGASIRYGVTRHHQGRWQARIGRVAGNKDLYLGTFTGQEEAAEA 352
DB 352 QEYVASLRKSGSFGSRGASIRYGVTRHHQGRWQARIGRVAGNKDLYLGTFTGQEEAAEA 411
QY 353 YDIAIKRGRGTAVTNFDMRNMYNKATLESPLIPGSAKELKE-----ANR 399
DB 412 YDVAIKRGRGSAVTNFMRSYDVKSLLESTLPIGGAKEKDKMEQVELRVENHRAQ 471
QY 400 PVPSPMMTISNNVSEBENSASQWAAVGHQGVLDLSLHQHQRERYNGYNGNLSSESA 459
DB 472 EDHSSIMNSHLTQGIINNYAAGTTATHHNNHNAFAHQPPQCTTIHYPGQRIN--- 527
QY 460 RACFKQEDDQ-----HHFLSNTQSL--MTNIDHQSYSVSDSVT 495
DB 528 --WCKQEQDNDASHLSYSIDHQLQGNNTHTNFFHTNSGLHPLMSMD---SASIDNSS 582
QY 496 VCGNVV--GYGQYQGF-AAPV-NCDAVAASEFDYN-----526
DB 583 SSNSVYDYGGGGGYNNVPIPGTITTVVANDQDQPNRSHGFGDNEIKALGYESVYSGTT 642
QY 527 -----ARNHYFAQQOQTQSPGDPFAAMTNVNVGSMYHYGGEGGEVAPT 572
DB 643 DPYHAERNYLYLTQQQPS-----VDAVKASAYDQGSACNTWVPT 683

RESULT 5

ABR40877
ID ABR40877 standard; protein; 692 AA.
AC ABR40877;
DT 16-MAY-2003 (first entry)
DE Oryza sativa oil trait related protein sequence SEQ ID NO:530.
KW plant; oil trait; oil phenotype; altered lipid profile; MAP kinase; receptor-like protein kinase; mitogen activated protein kinase; oil; LIP15-like transcription factor calcosin; ATP citrate lyase; SNF1; CKC-like transcription factor; antisense inhibition; co-suppression; transgenic plant.
XX
OS Oryza sativa.
XX WO2003002751-A2.
XX 09-JAN-2003.
XX 27-JUN-2002; 2002WO-US020152.
XX 29-JUN-2001; 2001US-0301913P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT, Jones TC, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B, Tarczynski MC;
XX WPI; 2003-201509/19.
XX Novel nucleotide fragment encoding polypeptides having receptor-like protein kinase activity, calcosin-like activity, useful for altering oil phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX Claim 12; Page 538-540; 542pp; English.
XX The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (PP) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, LIP15-like transcription factor activity, calcosin-like activity, ATP citrate lyase activity, SNF1-like activity, and CKC-like transcription factor activity. Also described: (I) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention

XX
SQ Sequence 692 AA;

Query Match 36.3%; Score 1131; DB 6; Length 692;
Best Local Similarity 40.1%; Pred. No. 2.7e-87;
Matches 278; Conservative 65; Mismatches 125; Indels 236; Gaps 23;

QY 3 NNWLGFSLSPYEQNHRRKDVYSSSTTTTVDVAGYCYDPTAASDESSAIOISFSPFGV 62
DB 5 NNWLAFSLSPQQLPP-----SQNSTLSIAAA-----TTTAGDSSGDCVCFNIP--- 50
QY 63 VDAFTDRNNHSRDWDINGCACNNTHNDEQGPKEFLGRTTIYNTNENVG-----117
DB 51 -----QDWSMRGSELSALVAE-----PKLEDFLGISFSEQHHGGKGVIPS 94
QY 118 SGSGCY-----GGG-----DG 128

```

Db 95 SAAACYSAGSSVGLYLPFPSSSSLOPADSVMTATSPVVAHDVSGGWSVAAAAAAS 154
QY 129 GGSGLGSMITKWLNRQPDVNDVQENQENNAAGLSLSMNS----- 168
Db 155 GNGIGLSMITKWLNRQFPAP-----QPAQALSLSMNMAGTTTAQGGGAMALLAGAGE 206
QY 169 -----STSC-----DNNNDNNNVAAQKTIIDS-----V 193
Db 207 RGRTPPASELSITSAGHATTATWAGGKEINEBGSAGAVAVGSESGSAGAVVEAGAA 266
QY 194 EATPKKTIESFGORTSYRGVTRHRTWGRVYEAHLWNSCKRGQTRKGRQVYLGVDKKE 253
Db 267 AAARKSVDFGORTSYRGVTRHRTWGRVYEAHLWNSCKRGQTRKGRQVYLGVDKKE 323
QY 254 KAARAYDLAALKYGGTTTTFNPFMSVEKEVEEMKMTROEYVASLRKSSGFSRGASTY 313
Db 324 KAARAYDLAALKYGGPTTTFNPFNNYKELEEMKMTROEYVASLRKSSGFSRGASTY 383
QY 314 RGVTRHHQGRWQARIGRVAGNKDLYLGTFGTDEEAAEYDIAAIFRGLTAVTNFDMNR 373
Db 384 RGVTRHHQGRWQARIGRVAGNKDLYLGTFGTDEEAAEYDIAAIFRGLTAVTNFDMNR 443
QY 374 YNVKAILSPSLGSAKELKANRPV-SMMNINNVNS-----ESENSASGW 421
Db 444 YVKSILDSALPVGTAARLKDAAEAAAAYDVGRVIAHLGGDGAAYAAHGHCHHHSAAA 503
QY 422 QNAAVQ-----HHQGVLDLSLHQHQRVYNGYVANGNLSSESARACFKQEDDQHFLS 474
Db 504 PTIAFQAAAAAPPHPAAG---LYHPYAQPLRGW-----C-KQEQD-HAVTA 543
QY 475 NTQSLTNIDHQSSVDDSVTVCGNVVGYGOGYGAAPVNCDAYAASEFDYNARNHYFA 534
Db 544 AHSLOQ-DLHH-----LNLGAAAAA-----HDFFS 567
QY 535 QQOQTQSGPGDPPAAMTNVNGSN-MYIHGEGG 567
Db 568 QAMQOQHGLSIDNASLEHSGSVVYNGDNGG 601

RESULT 6
ABR40771
ID ABR40771 standard; protein; 553 AA.
AC
XX ABR40771;
XX 16-MAY-2003 (first entry)
DE Glycine max oil trait related protein sequence SEQ ID NO:362.
KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
KW transgenic plant.
XX
OS Glycine max.
XX
XX WO2003002751-A2.
XX
XX 09-JAN-2003.
XX
XX 27-JUN-2002; 2002WO-US020152.
XX
XX 29-JUN-2001; 2001US-0301913P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
XX Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
XX Tarczynski NC;

```

```

DR WPI: 2003-201509/19.
DR N-PSDB; ACC00806.
XX
PT Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX
PS Claim 12; Page 382-383; 542pp; English.
XX
CC The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, Lip15-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity, and CKC-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV); and (5)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00806 to ACC00808 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 553 AA;
Query Match 35.9%; Score 1119.5; DB 6; Length 553;
Best Local Similarity 43.5%; Pred. No. 1.8e-86;
Matches 275; Conservative 77; Mismatches 147; Indels 133; Gaps 23;
QY 1 MNNWLGFSLPYEQNHHRKDVSSSTTTTVDVAGEYCYDPTAASDESSAQTGFPSPFG 60
Db 1 MNNWLSFPLS-----FTHSLPAHDLAQYHQFS 31
QY 61 VVDAFTDNNKSHRDWD-INGCACNNIHNDQDQPKLENFLGRTTTTYNTNENVDGSG 119
Db 32 IGLVNNNDNPFQNHDMNLINTHSNEI-----PKVADFLG-----VKSSENQSDLAA 79
QY 120 SGCYGGDGGGGSLGSLKTLNRPQVDNVN--QENGNA-AKGLSLSMNS--STSCDN 174
Db 80 LNEIHSNDSVLYFTNLSLV--MQNPLDTPSNYQENANSLQSLTSLMSGKSTCET 137
QY 175 NDSNNNVVAAOKTIDDSVEATPKKTIESTFCORTSYRGVTRHRTWGRVYEAHLWNSCKR 234
Db 138 SCENSTNT-----TVEVAPRRTLDTFCORTSYRGVTRHRTWGRVYEAHLWNSCKR 188
QY 235 EGOTRKGQVYLGVDKEEKAARAYDLAALKYGGTTTTFNPFMSVEKEVEEMKMTROE 294
Db 189 EGQSRKGRQVYLGVDKEEKAARAYDLAALKYGGTTTTFNPFMSVEKEVEEMKMTROE 248
QY 295 YVASLRKSSGFSRGASTYRGVTRHHQGRWQARIGRVAGNKDLYLGTFGTQEEAAEAYD 354
Db 249 FVAALRRKSSGFSRGASMYRGVTRHHQGRWQARIGRVAGNKDLYLGTFGTQEEAAEAYD 308
QY 355 IAAIKFRLTAVTNFDMNRYNVAKILSPSLPI-GSAARLKEA-----NRPVSMNMI- 407
Db 309 IAAIKFRLNATVTFNDFSRYDVKAILSENTPILGGAARLKEAQALESKRREBEMTALG 368
QY 408 SNNVSESENSASGWNAA---VQHH-----QGVDLSLH-QHQY---RYNGY 448
Db 369 SSSITFYGTSSASSRLHAYPLMQHHQFEPQPLLTQLNHDISSSHFSHQDPLHHQYI 428
QY 449 YNGNLSSESARACFK-QEDDQHH--FLSNQTSL---MTNIDHSSVS-----DSDVTVC 497
Db 429 QTQLQLHQOQSGASSYFQNNAAQFYNGYLQNHLPALQGMNMGSSSSSSSVLENNNSNN 488
QY 498 GNVVGYGQYQGFAPVNCDA-----YAASEFDYNARNHYFAQQOQTQOSPGD 546
Db 489 NNVGFGVG-SGFGMASNATAGTNTVGTAEELGLVKVDYDMPAGGCGGWSAASDMOTSGG- 546

```

QY 547 FPAAMTNNVGSNNYYHGBGGGEVAPTFTVMD 578

DB 547 -----VFTMND 553

RESULT 7

ID ABR40767
XX ABR40767 standard; protein; 486 AA.

AC ABR40767;

DT 16-MAY-2003 (first entry)

DE Glycine max oil trait related protein sequence SEQ ID NO:354.

XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
transgenic plant.

XX Glycine max.

XX WO2003002751-A2.

PN 09-JAN-2003.

PD 27-JUN-2002; 2002WO-US020152.

PF 29-JUN-2001; 2001US-0301913P.

PR (DUPO) DU PONT DE NEMOURS & CO E I.

PA (PION-) PIONEER HI-BRED INT INC.

XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT,
PI Jones TJ, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H, Shen B;
PI Tarczynski MC;

XX WPI; 2003-201509/19.

DR N-PSDB; ACC00802.

XX Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.

PS Claim 12; Page 372-373; 542pp; English.

XX The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (pp) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, Lip15-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity and CKC-like transcription factor activity. Also described: (I)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
CC invention

SQ Sequence 486 AA;

Query Match 34.5%; Score 1075.5; DB 6; Length 486;
Best Local Similarity 50.2%; Pred. NO. 8.7e-83;
Matches 240; Conservative 45; Mismatches 99; Indels 95; Gaps 12;

QY 167 NSTSTCDNNNDNNVAGCTIDSVETAPKKTIESFGORTSIYRGVTRHRTGRVEAH 226

DB 4 NGESSSDNKQP--NTSAALDSTQTCAIETAPRKSIDTFTGORTSIYRGVTRHRTGRVEAH 61
QY 227 LWDNSCKREGOTRKGROYVYLGIDYKKEKAAAYDLAALKYAGTTTNTNPFMSEYEKEVEE 286
DB 62 LWDNSCRREGQTRKGROYVYLGIDYKKEKAAAYDLAALKYAGTTTNTNPFISHYEKELEE 121
QY 287 MKHMTROEYVASLRKSSGFGSGASIRGVTRHHQHGWRQARIQVAGNKDLYLTGTGTQ 346
DB 122 MKHMTROEYVASLRKSSGFGSGASIRGVTRHHQHGWRQARIQVAGNKDLYLTGTGTQ 181
QY 347 EBAEAYDIAIKPRGLTAVTNFDMNRYNVKALLESPLPIGSAAKRLKE----- 396
DB 182 EBAEAYDVAAIKFRGLSAVTFNMSRYDVKSIIESTITPIGGAARLQMEQVELSDVN 241
QY 397 ---ANRPVPSMMIISNVSESENASGWCNAAVOHQGVDSLHLHQHQRNGYNGYNGN 453
DB 242 GHRADQVDHSIIIMSSHLTQGINNVAGGTAT--HNNWNAHAFQPOCTTTHYPYQOR 299
QY 454 LSSESARACFKQEDDQ-----HHFLSNTQSL--MTNIDHQSSV 489
DB 300 IN-----WCKQEQDQNSDAPHSLSYSDIHQLQLGNNGTHNFTNSGLHPMLSMD---SA 351
QY 490 SDDSVTVCGNVV--GYGGYQGF-AAPV--NCDAYAASEFDYN----- 526
DB 352 STDNSSSSNVYDGYGGGGYVWPMGT--TAVVASDGDQNPNSNHGFGDNEIKALGYES 411
QY 527 -----ARNHYFAOQOQTCQSPGDFPAAMTNNVGSNNYYHGBGGGEVAPT 572
DB 412 VYGSATDSYHAHARNLYLTQOQ-----SSVDTVKASAYDQGSACNTWVPT 458

RESULT 8

ID ADD30879 standard; protein; 574 AA.

XX AC ADD30879;

DT 15-JAN-2004 (first entry)

DE Plant yield-related protein from clone G1793.

XX transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.

XX Arabidopsis thaliana.

XX WO2003013227-A2.

XX 20-FEB-2003.

XX 09-AUG-2002; 2002WO-US025805.

XX 09-AUG-2001; 2001US-0310847P.

XX 19-NOV-2001; 2001US-0336049P.

XX 11-DEC-2001; 2001US-0338692P.

XX 14-JUN-2002; 2002US-00171468.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;

XX Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;

XX Brown PE;

XX WPI; 2003-248221/24.

XX N-PSDB; ADD30878.

XX New plant transcription factor polynucleotides and polypeptides, useful

PT in producing transgenic plants with commercially valuable properties,

PT such as an alteration in a plant growth characteristic, e.g. growth rate

XX or apomixis.

XX Disclosure; SEQ ID NO 908; 454pp; English.

AC ABR40836;
 XX DT 16-MAY-2003 (first entry)
 XX DE Arabidopsis thaliana oil trait related protein sequence SEQ ID NO:428.
 DE plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNPL;
 KW CXC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant.
 XX OS Arabidopsis thaliana.
 XX PN WO2003002751-A2.
 XX PD 09-JAN-2003.
 XX PF 27-JUN-2002; 2002WO-US020152.
 XX PR 29-JUN-2001; 2001US-0301913P.
 XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PA (PION-) PIONEER HI-BRED INT INC.
 XX PI Allen SM, Allen WB, Cahoon RE, Epeibaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX WPI; 2003-201509/19.
 XX DT Novel nucleotide fragment encoding polypeptides having receptor-like
 XX protein kinase activity, caleosin-like activity, useful for altering oil
 XX phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 12; Page 469-471; 542pp; English.
 XX The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNPL-like
 CC activity and CXC-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (II) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX SQ Sequence 516 AA;
 Query Match 33.4%; Score 1040.5; DB 6; Length 516;
 Best Local Similarity 44.6%; Pred. No. 9.3e-80;
 Matches 254; Conservative 70; Mismatches 114; Indels 131; Gaps 22;
 86 NIHNDEODG---PKLENFLGRITTT-IYNTNEN-----VGDGSGSGCYGGGGGSLGL 135
 3 NPHGGGEGGEVPKVAFLGVSKSGDHTDNLVYNDIHQTNASDYFOTN---SL-L 57
 136 SMITWLRNPQVNDVNOENQNAAGLSLNNSS-----TSCDNNDSN 179
 58 PTVTVCASNAP-NNYELQESAHNLQSLTSLMGSTGAAAEVATVKASPAETSDNSSST 116
 180 NNVAQKTIIDSVEATPKTIESFGORTSIYRGVTRHRTWGRYEAHLWNSCKREGQTR 239
 117 N---TSGAI---VEATPRILEIFGORTSIYRGVTRHRTWGRYEAHLWNSCKREGQSR 170

QY 240 KGRVYLGDDYKKEKKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMTQEVVASL 299
 DB 171 KGRQ---GGYDKEKKAARAYDLAALKYWGPSTTTTTFITNYEKEVEEMKMTQEVVASI 227
 QY 300 RKSSGSRGASIVRGVTRHGHGRWQARIGRVAGNKDLYLGTGTCQEEAAEAYDIAAIK 359
 DB 228 RKSSGSRGASIVRGVTRHGHGRWQARIGRVAGNKDLYLGTGTCQEEAAEAYDIAAIK 287
 QY 360 FRGLTAVTNFDMNRVNVKAILSPSLPI-GSAAKRLKEA-----NRVPVSMNMISNNVSE 413
 DB 288 FRGLNAVNTFENRYNDYKAILSENTLPIGGAARKLKEAQALESRRREEMIALGNSFHQ 347
 QY 414 SENSASGQNAA-----VOHHQGV-----DLS----- 435
 DB 348 -YGAASGSSSVASSRLQLQFPLSIQQPFELHHQPLLTQLQNNNDISQVHDSFSYIOT 406
 QY 436 LLHGHQERYNGYVYNGGNLSSSARACFKQEDDOHFLSNTQSLMTNIDHQSSVSDSVT 495
 DB 407 QLHLHQOQTNNYLOSSTLSQ-----YNAVLSQNPGLL-----HGFVSDNN-- 448
 QY 496 VCGNVVGYGGYQGPAAFPVNCDAVAASEDFYNARNHYFAOQQQTQSPGQDFPAAMTNV 555
 DB 449 ---NTSGFLGNNIGIGSSSTVGSSAEEFPAVKVDY-----DMPPS----- 487
 QY 556 GSNMYVHCEGGGVAP-----TFTVWND 578
 DB 488 CGATGYGWNESGESAQSGNPGVFTMNE 516
 RESULT 12
 ABB79640
 ID ABB79640 standard; protein; 585 AA.
 AC ABB79640;
 XX DT 21-OCT-2002 (first entry)
 XX Cotton AINTEGUMENTA-like polypeptide GhANT1.
 DE AINTEGUMENTA; ANT-like polypeptide; GhANT1; transgenic plant; cotton;
 KW plant.
 XX Gossypium hirsutum.
 XX WO200259332-A2.
 XX 01-AUG-2002.
 XX 19-DEC-2001; 2001WO-US049294.
 XX 21-DEC-2000; 2000US-0257896P.
 XX (MONS) MONSANTO TECHNOLOGY LLC.
 XX He SS, Dotson SB;
 XX WPI; 2002-599798/64.
 XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in
 PT improving agronomic, horticultural, and quality traits of plants, such as
 PT increased size of plant organs.
 XX Claim 3; Page 158-161; 169pp; English.
 XX The present sequence is the protein sequence of GhANT1, a newly
 CC identified AINTEGUMENTA-like (ANT-like) polypeptide of cotton. The
 CC sequence was predicted from a cDNA clone (see ABN84485) isolated in a
 CC cotton DNA database screening. The invention provides nucleic acids
 CC encoding ANT-like polypeptides comprising, in N-terminal to C-terminal
 CC direction, 2 AP2 DNA binding domains followed by an amino acid
 CC subsequence selected from those given in ABB79629-35. ANT-like
 CC polypeptides were identified in soybean, rice, cotton and corn (see
 CC ABB79636-41). Nucleic acids (see ABN84480-86) encoding the ANT-like

Qy 292 ROEYVASLRKSSGFSRGASIRYGVTRHHQGRQWQARIGRVAGNKDLYLGTFTQTEAAE 351
Db 259 ROEYVASLRKSSGFSRGASIRYGVTRHHQGRQWQARIGRVAGNKDLYLGTFTQTEAAE 318
Qy 352 AYDIAAIKFRGLTAVTNFDMRNVNKALLESPLPSIGSAARLKEANRPVPSMMISNVV 411
Db 319 AYDIAAIKFRGLTAVTNFDMRNVNKALLESPLPSIGSAARLKEANRPVPSMMISNVV 371
Qy 412 SESENSASG-----WQNAAVQHQGVLDLSL-----LH 438
Db 372 TDEERHSGSGCMFTFSQPAASSTVGFAINIKODPSDYNVNLGYNDTGASLSAXNPFI 431
Qy 439 QHBERYNGYY-----NGCN--LSESARAFK-----QEDDQHFLFNTQSMT-- 481
Db 432 HHSFNGSGFOSPGFCTGNGGNSAVGSIINGLFNRGYGVQEQEASSAATNSHPLATPI 491
Qy 482 --NIDHCSSVSDSVTCVGNVYGG 505
Db 492 VLNNSDYESSSS-----GYGG 508

RESULT 14
ABR40853
ID ABR40853 standard; protein; 470 AA.
AC ABR40853;
XX
XX 16-MAY-2003 (first entry)
XX
XX Catalpa sp oil trait related protein sequence SEQ ID NO:485.
DE
DE Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
KW transgenic plant.
XX
XX Catalpa sp.
XX
XX W02003002751-A2.
XX
XX 09-JAN-2003.
XX
XX 27-JUN-2002; 2002WO-US020152.
XX
XX 29-JUN-2001; 2001US-0301913P.
XX
XX (DUPO) DU FONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
XX Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
XX Tarczyński MC;
XX
XX WFI; 2003-201509/19.
XX N-PSDS; ACC00847.
XX
XX Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX
XX Claim 12; Page 501-503; 542pp; English.
XX
XX The present invention describes an isolated nucleotide fragment (I)
XX comprising a nucleic acid sequence (NS) chosen from a NS encoding a
XX polypeptide (PP) having receptor-like protein kinase activity, mitogen
XX activated protein (MAP)-kinase activity, Lip15-like transcription factor
XX activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
XX activity and CKC-like transcription factor activity. Also described: (1)
XX complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
XX (II), operably linked to a regulatory sequence; (3) a plant (IV)
XX comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)

CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC created transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00826 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
XX invention
XX
SQ Sequence 470 AA;
Query Match 32.2%; Score 1003; DB 6; Length 470;
Best Local Similarity 45.7%; Pred. No. 1.3e-76;
Matches 235; Conservative 49; Mismatches 94; Indels 136; Gaps 16;
Qy 2 NNNWLGFLSLPYEQNHHRKDVYSSSTTTTVDVAGYCYDPTAASDESSAIQTSPSPGCV 61
Db 7 HQWLAFLSL-----NHS-----LFHPSFNSPSSC-----GV 35
Qy 62 VYDAFTRNNSHSRDWDINGCACNNIHD--EQDGPKELENFLGRTRTTTYNTNENVGDSGG 119
Db 36 V-----NQENG-----AGGAATDHLVLASGGPKLEDFLGSC-----GSS 73
Qy 120 SCYGGGSGGGSLGLSMIKTWLRNQPDVNDVQNGNAKGLSLSMNSSTSCDNNDSN 179
Db 74 GCHTQYDGGAGGVNISDVEIF-----GSELKTIAASFUGLSHERSDSQ 120
Qy 180 NNVAQGTIDDSVEATPKTIESFGQRTSIYRGVTRHRTGRYEAHLWDNSCKEGQTR 239
Db 121 KQTAVAAPP-----EPKAKAAETFGQRTSIFRGVTRHRTGRYEAHLWDNSCREGQSR 175
Qy 240 KGRQVYLGSDYKKEKAAAYDLAALKYGTGTTTTFNPFNSEYKEVEEMKHTROEYVSL 299
Db 176 KGRQVYLGSDYKKEKAAAYDLAALKYGTGTTTTFNPFNSEYKEVEEMKHTROEYVSL 235
Qy 300 RKSSGFSRGASIRYGVTRHHQGRQWQARIGRVAGNKDLYLGTFTQTEAAEAYDIAIK 359
Db 236 RKSSGFSRGASIRYGVTRHHQGRQWQARIGRVAGNKDLYLGTFTQTEAAEAYDIAIK 295
Qy 360 FRGLTAVTNFDMRNVNKALLESPLPSIGSAARLKEANRPV-----SMMISNVV 411
Db 296 FRGLTAVTNFDMRNVNKALLESPLPSIGSAARLKEANRPV-----SMMISNVV 354
Qy 412 S-ESENSAS-----GWCNAVQH-----HQGVLDLSLH----- 438
Db 355 SPPSPNSAAINFALPIKQDHSSTLWNLGYQNSTITNTNRPNTSVSLFHDNSFMSMEFQ 414
Qy 439 -----QHBERYNGYYNGNLSES 458
Db 415 ANNNEGFFNGCGYFOHQE-----NGTNSSTS 441

RESULT 15
ABR40852
ID ABR40852 standard; protein; 512 AA.
XX
XX ABR40852;
XX
XX 16-MAY-2003 (first entry)
XX
XX Glycine max oil trait related protein sequence SEQ ID NO:483.
XX
XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
KW transgenic plant.
XX
XX Glycine max.
XX W02003002751-A2.
XX 09-JAN-2003.
XX

XX 27-JUN-2002; 2002MO-US020152.
XX PF (DUPO) DU PONT DE NEMOURS & CO E I.
XX PR (PION-) PIONEER HI-BRED INT INC.
XX PA
XX PA
XX PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
PI Tarczyński MC;
XX WPI; 2003-201509/19.
DR N-PSDB; ACC00846.
XX Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, calesiosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX
PS Claim 12; Page 499-500; 542pp; English.
XX
CC The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, Lipid-like transcription factor
CC activity, calesiosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity, and CKC-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00826 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 512 AA;
Query Match 32.0%; Score 999; DB 6; Length 512;
Best Local Similarity 45.9%; Pred. No. 3.2e-76;
Matches 235; Conservative 61; Mismatches 126; Indels 90; Gaps 18;
QY 2 NNNWLGSLSPYEONHHRKDYSSTTTTVDVAGEYCYDPTAASDESS---AQTSPFSP 58
DB 5 STNWLSPSLSP-----MDMLRTPPEQFVQ-----YD--AASDTSSHHYLDNLYTNG 49
QY 59 FGWVVDATRDNNSHRDWDINGCACNNIHDEQD----GPKLEFLGRITTIYNTNEN- 113
DB 50 WG-----NGSLKFEQNLNHSDFVQSSSSVSHAPPKLEDFLGDSAVMRYSDSQ 100
QY 114 --VGDSG-----GSGCYGGGGGG--GSLGLSMIKTWRNQPVNDVNDQENG 156
DB 101 TETQSSSLTHYDHHHHHHGSSAYFGGDHQLKAITGTFQAFST-----NSGSEVDDSGASI 156
QY 157 NAAKGLSLMSNSTSCDN-----NNDSSNNVVQAQKTIIDDSVEAT-----PKYTIES 203
DB 157 GRAQGSSEFQTHIESSVNEFAAFSGGTNTGGTSLSLAVAQSEKAVAAAEASDRSKKVVDT 216
QY 204 FGORTSIYRGVTRHWTGRYEHLWDNSCKREGQKRGQVYLGKYDKEEKAARAYDLAA 263
DB 217 FGORTSIYRGVTRHWTGRYEHLWDNSCRREGQKRGQVYLGKYDKEEKAARSYDLAA 276
QY 264 LKYWGTTTTNTPMSEYKEVEEMKHMTRQBYVASLRKXSGSGFRGASIRGVTRHHQHG 323
DB 277 LKYWGPTATTNPFVSNYSKEVEEMKHVTQKEFIASLRKXSGSGFRGASIRGVTRHHQHG 336
QY 324 RWQARIGRVAGNKDLYLGTFTGTEEAABAYDIAIKFRGLTAVTNFDMRYNKAILESP 383
DB 337 RWQARIGRVAGNKDLYLGTFTGTEEAABAYDIAIKFRGANAVTNFENRYDVEAIMKS- 395

QY 384 SLPIGSAARLK-----EANRPVPSMMWISNNVSESENSASGWMONAAVQH---HQGV-- 432
DB 396 SLIPVGAARLKLSLESEKALPVSS-----SSSSSQOQNPOCGNVSAINFSSIHQPIAS 451
QY 433 -----DLSLLHQHQRNYGY-YNGGNLSS 456
DB 452 IPCGIPFDSSTAYVHHNLFQHFHPTNAGTAAS 483

Search completed: July 15, 2004, 10:01:52
Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 10:00:59 ; Search time 19 Seconds
(without alignments)
1573.234 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119

Sequence: 1 MNNNMLGFLSPYEQNHHRK.....YYHGEGGGEVAPFTTWNNDN 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:**

- 1: /cgn2_6/prodata/2/iaa/5A COMB pep.*
- 2: /cgn2_6/prodata/2/iaa/5B COMB pep.*
- 3: /cgn2_6/prodata/2/iaa/6A COMB pep.*
- 4: /cgn2_6/prodata/2/iaa/6B COMB pep.*
- 5: /cgn2_6/prodata/2/iaa/PTUS COMB pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	913.5	29.3	548	4	US-09-227-421-5
2	913.5	29.3	548	4	US-09-479-855-5
3	891	28.6	555	4	US-09-227-421-2
4	891	28.6	555	4	US-09-479-855-2
5	461.5	14.8	432	2	US-08-700-152A-4
6	345	11.1	77	3	US-08-912-272-9
7	345	11.1	77	4	US-09-026-039-9
8	328	10.5	69	3	US-08-912-272-10
9	328	10.5	69	4	US-09-026-039-10
10	188	6.0	67	3	US-08-912-272-11
11	188	6.0	67	4	US-09-026-039-11
12	181	5.8	67	2	US-08-700-152A-1
13	176.5	5.7	68	2	US-08-700-152A-2
14	176.5	5.7	68	3	US-08-912-272-5
15	176.5	5.7	68	4	US-09-026-039-5
16	174	5.6	67	3	US-08-912-272-4
17	174	5.6	67	4	US-09-026-039-4
18	152.5	4.9	375	4	US-09-533-029-12
19	151	4.8	328	4	US-09-964-850-4
20	150	4.8	171	4	US-09-533-029-82
21	150	4.8	248	4	US-09-533-029-22
22	149.5	4.8	328	4	US-09-964-850-6
23	145	4.6	248	4	US-09-202-161B-3
24	144.5	4.6	306	4	US-09-964-850-2
25	141	4.5	264	3	US-09-964-731-4
26	135	4.3	216	4	US-08-894-731-1
27	135	4.3	216	4	US-09-198-119C-15
					US-09-301-666A-2
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 4, Appli
					Sequence 9, Appli
					Sequence 10, Appli
					Sequence 10, Appli
					Sequence 11, Appli
					Sequence 11, Appli
					Sequence 1, Appli
					Sequence 2, Appli
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 4, Appli
					Sequence 12, Appli
					Sequence 4, Appli
					Sequence 82, Appli
					Sequence 22, Appli
					Sequence 6, Appli
					Sequence 3, Appli
					Sequence 2, Appli
					Sequence 4, Appli
					Sequence 15, Appli
					Sequence 2, Appli

28 135 4.3 216 4 US-09-301-217-2 Sequence 2, Appli
29 134.5 4.3 188 4 US-09-198-119C-65 Sequence 65, Appli
30 133 4.3 66 3 US-08-912-272-29 Sequence 29, Appli
31 133 4.3 66 4 US-09-026-039-29 Sequence 29, Appli
32 132.5 4.2 344 4 US-09-533-029-66 Sequence 66, Appli
33 132 4.2 243 4 US-09-533-029-34 Sequence 34, Appli
34 129.5 4.2 68 3 US-08-912-272-28 Sequence 28, Appli
35 129.5 4.2 68 4 US-09-026-039-28 Sequence 28, Appli
36 129 4.1 328 3 US-09-300-672-2 Sequence 2, Appli
37 128.5 4.1 58 4 US-09-202-161B-27 Sequence 2, Appli
38 127.5 4.1 161 4 US-09-198-119C-13 Sequence 13, Appli
39 126.5 4.1 216 4 US-09-301-666A-8 Sequence 8, Appli
40 126.5 4.1 216 4 US-09-301-217-8 Sequence 8, Appli
41 126.5 4.1 115 4 US-09-198-119C-95 Sequence 95, Appli
42 125 4.0 640 4 US-09-336-115C-14 Sequence 14, Appli
43 124.5 4.0 213 2 US-08-949-603-2 Sequence 2, Appli
44 124 4.0 213 2 US-08-706-270A-2 Sequence 2, Appli
45 124 4.0 213 2 US-08-706-270A-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-227-421-5
; Sequence 5, Application US/09227421
; Patent No. 6559357
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
; FILE REFERENCE: 023070-090700PC
; CURRENT APPLICATION NUMBER: US/09/227,421
; CURRENT FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 09/227,421
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: PRT
; LENGTH: 548
; ORGANISM: Brassica napus
US-09-227-421-5

Query Match 29.3%; Score 913.5; DB 4; Length 548;

Best Local Similarity 42.0%; Pred. No. 8e-74; Indels 125; Gaps 13;

Matches 216; Conservative 57; Mismatches 116; Indels 125; Gaps 13;

QY 4 NMLGFLSP-YEQNHHRKDVYSSSTTTVV-----DVAGEY--CYDPTAASDESS 49
DB 14 NLGFLSSNMLKMGGEALYSSSSSVATSSVPPQLVVDNNSNYGVYGSNLAAREMY 73
QY 50 ALOTGFPFPFG---VVDAFTDNNHSHRDWINGCANNIHNDRQDQKLENFLG---- 102
DB 74 SQMVSPLRSDGSLCLMEALNR--SSHS-----NNHHSQVSPFKMEDFFGTHH 121
QY 103 -----RTTIYNTNENVGDSGCGYGGGGGSLGSLMIKTLRNQ 145
DB 122 NTSKHEAMDLSLDFYNTHAPNNTNFOE-----FFSPQT--RNH 162
QY 146 PVDVNDNQEN-----GNAAGLSLSMN--SSTCDNNDSNNNVVAGK 187
DB 163 HEEETRNVEDPGLTHGGGSFNVGVGFQQSLSLSPGSSCITASHHHQNTQNHQ 222
QY 188 TDDSVETAP-----KKTIESFGORTSIYRGVTRH 217
DB 223 QISEALVETSAQFETTMAAAAKKRGQEVVVGQKQIVHRKSIDTFCORTSIYRGVTRH 282
QY 218 RMTGRYEHLWDSCKREGQTRKGRQVYLGGLGYDKEEKAARAYDLAALYKWTGTTTTFPM 277
DB 283 RMTGRYEHLWDSCKREGQTRKGRQVYLGGLGYDMEKAARAYDLAALYKWTGTTTTFPM 342


```

; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,152A
; FILING DATE: 20-AUG-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067200US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-700-152A-4

Query Match 14.8%; Score 461.5; DB 2; Length 432;
Best Local Similarity 34.3%; Pred. No. 3.4e-33;
Matches 114; Conservative 59; Mismatches 112; Indels 47; Gaps 8

Qy 198 KKTIESFGQRTSYRGVTRHWTRGRVEAHLDWNSCKREGQTRKGRQVYLGGYCKEEKAAAR 257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 KGRRGPRSRSSQVRGTFYRRTGRWESHWD--C-----GKVYVGGFDTAHAAR 168

Qy 258 AYDLAALKYWGTTTTFNFPNSEYEKEVEEMKHWTRQEVASLRKSKSGFSRGASTYRGVT 317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 AYDRAAIKFGVRADINFNIDDDDDLUQNTLTKEEFVHLRQSTGTFPRGSKYRGVT 228

Qy 318 RHQHGRQWQARIQIRVAGNKDLYLGTQGTQEEAAEAYDIAAKFEGLTAVTNFDMNRNVK 377
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 L-HKCGWEARMGQFLGKVKYVILGLPDTVEAARAYDKAAIKCKGKDAVTNFDPSIYDEE 287

Qy 378 AIUESPLPGSAAKRLKEANRPVPSMMMLSNNVSESENSASGQWNAVQHHQ----- 430
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 LNAES-----SGNPTTPDHNLDLSLGNANSKHKHSQDRLRMNQOQDLSLH 334

Qy 431 -----GVDLGLLHQHQBRYNGYVNGNLSESARACFKQEDDQHHF--LSNTQSLMTNI 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 SNEVLGLGQTGMLNHTNSNHQPPGSSNIGSGGFSLPFAEN-HRFDGRASTNQVLTA 393

Qy 484 -----DHQSSVSDSVTVCGNVVGYGGYQ 507
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 AASGGSFPHHNQIFNSTSTPHQWLOTNGFQ 425

RESULT 6
US-08-912-272-9
; Sequence 9, Application US/08912272
; Patent No. 6093874
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP

```

STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/912,272
FILING DATE: 15-AUG-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/879,827
FILING DATE: 20-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/700,152
FILING DATE: 20-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067220US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..77
OTHER INFORMATION: /note= "ANT-R1 direct repeat"

NAME/KEY: Region
LOCATION: 44..59
OTHER INFORMATION: /note= "putative ANT-R1 amphipathic
OTHER INFORMATION: alpha-helix (SEQ ID NO:37)"

Query Match 11.1%; Score 345; DB 3; Length 77;
Best Local Similarity 80.5%; Pred. No. 8.3e-24;
Matches 62; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 208 TSIYRGVTRHRTWGRYEAHLWDNSCKREGQTRKGRQVYLGQYDKEEKAARAYDLAALKYW 267

Db 1 TSQYRGVTRHRTWGRYEAHLWDNSFKEGHSRKGQVYLGQYDMEEKAARAYDLAALKYW 60

QY 268 GTTTTTFPMSEYKEV 284

Db 61 GPSTHTNFSAEYQKEI 77

RESULT 7

US-09-026-039-9

Sequence 9, Application US/09026039

Patent No. 6329567

GENERAL INFORMATION:

APPLICANT: Jofuku, K. Diane

APPLICANT: Okamuro, Jack K.

TITLE OF INVENTION: Methods for Improving Seeds

NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/912,272
FILING DATE: 15-AUG-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/879,827
FILING DATE: 20-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/700,152
FILING DATE: 20-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067220US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..77
OTHER INFORMATION: /note= "ANT-R1 direct repeat"

NAME/KEY: Region
LOCATION: 44..59
OTHER INFORMATION: /note= "putative ANT-R1 amphipathic
OTHER INFORMATION: alpha-helix (SEQ ID NO:37)"

Query Match 11.1%; Score 345; DB 3; Length 77;
Best Local Similarity 80.5%; Pred. No. 8.3e-24;
Matches 62; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 208 TSIYRGVTRHRTWGRYEAHLWDNSCKREGQTRKGRQVYLGQYDKEEKAARAYDLAALKYW 267

Db 1 TSQYRGVTRHRTWGRYEAHLWDNSFKEGHSRKGQVYLGQYDMEEKAARAYDLAALKYW 60

QY 268 GTTTTTFPMSEYKEV 284

Db 61 GPSTHTNFSAEYQKEI 77

RESULT 7

US-09-026-039-9

Sequence 9, Application US/09026039

Patent No. 6329567

GENERAL INFORMATION:

APPLICANT: Jofuku, K. Diane

APPLICANT: Okamuro, Jack K.

TITLE OF INVENTION: Methods for Improving Seeds

NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834


```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,272
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,272
; FILING DATE: 15-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/879,827
; FILING DATE: 20-JUN-1997
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067220US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..69
; OTHER INFORMATION: /note= "ANT-R2 direct repeat"
;
; NAME/KEY: Region
; LOCATION: 37..51
; OTHER INFORMATION: /note= "putative ANT-R2 amphipathic
; OTHER INFORMATION: alpha-helix"
;
; US-08-912-272-10

```

```

Query Match          10.5%; Score 328; DB 3; Length 69;
Best Local Similarity 91.0%; Pred. No. 2.4e-22;
Matches 61; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 310 ASIYGVTRHHQGRWQARIGRVAGNKDLYLGTGFTQEEAAAYDIAAIKFRGLTAVTNF 369
Db 1 ASIYGVTRHHQGRWQARIGRVAGNKDLYLGTGFTQEEAAAYDIAAIKFRGTNAVTF 60

QY 370 DMRNVV 376
Db 61 DITRYDV 67

```

```

RESULT 9
US-09-026-039-10
; Sequence 10, Application US/09026039
; Patent No. 6329567
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,039
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,272
; FILING DATE: 15-AUG-1997
; APPLICATION NUMBER: US 08/879,827
; FILING DATE: 20-JUN-1997
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..69
; OTHER INFORMATION: /note= "ANT-R2 direct repeat"
;
; NAME/KEY: Region
; LOCATION: 37..51
; OTHER INFORMATION: /note= "putative ANT-R2 amphipathic
; OTHER INFORMATION: alpha-helix"
;
; US-09-026-039-10

```

```

Query Match          10.5%; Score 328; DB 4; Length 69;
Best Local Similarity 91.0%; Pred. No. 2.4e-22;
Matches 61; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 310 ASIYGVTRHHQGRWQARIGRVAGNKDLYLGTGFTQEEAAAYDIAAIKFRGLTAVTNF 369
Db 1 ASIYGVTRHHQGRWQARIGRVAGNKDLYLGTGFTQEEAAAYDIAAIKFRGTNAVTF 60

QY 370 DMRNVV 376
Db 61 DITRYDV 67

```

```

RESULT 10
US-08-912-272-11
; Sequence 11, Application US/08912272
; Patent No. 6093874
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```



```

: NAME: Bastian, Kevin L.
: REGISTRATION NUMBER: 34, 774
: REFERENCE/DOCKET NUMBER: 023070-067200US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 67 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FEATURE:
: NAME/KEY: Domain
: LOCATION: 1..67
: OTHER INFORMATION: /note= "AP2TALA 2 (AP2) of Arabidopsis
: OTHER INFORMATION: direct repeat domain AP2-R1 consisting
: OTHER INFORMATION: of amino acids 129-195 of the AP2
: OTHER INFORMATION: protein"
:
US-08-700-152A-1

Query Match          5.8%; Score 181; DB 2; Length 67;
Best Local Similarity 46.8%; Pred. No. 4.2e-09;
Matches 36; Conservative 14; Mismatches 17; Indels 10; Gaps 2;

Qy      208 TSIYRGVTRHRTWTKRYEBAHLWDNCSCKREGQTRKGRQVVLGGYDKEEKAARAYDLAALKYW 267
      : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db      1 SSYRGVTFYFTRTGRWESHWD--C-----GKQVILGGFDTAHRAARAAYDRAAIKFR 50
      : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

Qy      268 GTTTTITNPPMSEYEKEY 284
      : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db      51 GVEADINFDYDDDL 67
      : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

```

```

RESULT 13
US-08-700-152A-2
; Sequence 2, Application US/08700152A
; Patent No. 5994622
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,152A
; FILING DATE: 20-AUG-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067200US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 1..68
; OTHER INFORMATION: /note= "APETALA 2 (AP2) of Arabidopsis
; OTHER INFORMATION: direct repeat domain AP2-R2 consisting
; OTHER INFORMATION: of amino acids 221-288 of the AP2
; OTHER INFORMATION: protein"
US-08-700-152A-2

Query Match 5.7%; Score 176.5; DB 2; Length 68;
Best Local Similarity 57.6%; Pred. No. 1.1e-08;
Matches 38; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 310 ASYRGVTRHHQGRQCARIGRVAGNKKDLVLTGFTQEEAAAYDIAAIKFRGLTAVTNF 369
Db 1 SSKIRGVTL-HKCRWEARMGQFLGKKYVYGLFDTEVEAARYDKAAIKNGKDAVTNF 59

QY 370 DMNRYN 375
Db 60 DPSIYD 65

RESULT 14
US-08-912-272-5
; Sequence 5, Application US/08912272
; Patent No. 6093874
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,272
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/879,827
; FILING DATE: 20-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067220US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..68
; OTHER INFORMATION: /note= "AP2-R2 direct repeat at
; OTHER INFORMATION: positions 221 to 288"

```

FEATURE:
NAME/KEY: Region
LOCATION: 33.50
OTHER INFORMATION: /note= "putative AP2-R2 amphipathic
OTHER INFORMATION: alpha-helix (SEQ ID NO:7)"
US-08-912-272-5

Query Match 5.7%; Score 176.5; DB 3; Length 68;
Best Local Similarity 57.6%; Pred. No. 1.1e-08;
Matches 38; Conservative 9; Mismatches 18; Indels 1; Gaps 1;
QY 310 ASIRGVTRHHQHGRVAGNCKDLYLGTGTGTEAAEAYDIAAIKFRGLTAVTNF 369
DB 1 SSKYRGVTL-HKCGRWEARMGQFLGKRYVYLGFLDTEVEAARAYDKAAIKNGKDAVTNF 59
QY 370 DMNRYN 375
DB 60 DPSIYD 65

RESULT 15
US-09-026-039-5
Sequence 5, Application US/09026039
Patent No. 6329567
GENERAL INFORMATION:
APPLICANT: Jofuku, K. Diane
APPLICANT: Okamuro, Jack K.
TITLE OF INVENTION: Methods for Improving Seeds
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/026.039
FILING DATE: 19-FEB-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,272
FILING DATE: 15-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/879,827
FILING DATE: 20-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/700,152
FILING DATE: 20-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..68
OTHER INFORMATION: /note= "AP2-R2 direct repeat at
positions 221 to 288"

FEATURE:
NAME/KEY: Region
LOCATION: 33.50
OTHER INFORMATION: /note= "putative AP2-R2 amphipathic
OTHER INFORMATION: alpha-helix (SEQ ID NO:7)"
US-09-026-039-5

Query Match 5.7%; Score 176.5; DB 4; Length 68;
Best Local Similarity 57.6%; Pred. No. 1.1e-08;
Matches 38; Conservative 9; Mismatches 18; Indels 1; Gaps 1;
QY 310 ASIRGVTRHHQHGRVAGNCKDLYLGTGTGTEAAEAYDIAAIKFRGLTAVTNF 369
DB 1 SSKYRGVTL-HKCGRWEARMGQFLGKRYVYLGFLDTEVEAARAYDKAAIKNGKDAVTNF 59
QY 370 DMNRYN 375
DB 60 DPSIYD 65

Search completed: July 15, 2004, 10:04:10
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 10:03:15 ; Search time 49 Seconds
(without alignments)
3693.319 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119

Sequence: 1 MNNMLGSLSPYEQNHK.....YVHGEGGEVAPTFTWMDN 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 1285345

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/1/pubaa/US07_PUBCOMB.pap:*
- 2: /cgn2_6/prodata/1/pubaa/PCT_NEW_PUB.pap:*
- 3: /cgn2_6/prodata/1/pubaa/US06_NEW_PUB.pap:*
- 4: /cgn2_6/prodata/1/pubaa/US06_PUBCOMB.pap:*
- 5: /cgn2_6/prodata/1/pubaa/US07_NEW_PUB.pap:*
- 6: /cgn2_6/prodata/1/pubaa/PCTUS_PUBCOMB.pap:*
- 7: /cgn2_6/prodata/1/pubaa/US08_NEW_PUB.pap:*
- 8: /cgn2_6/prodata/1/pubaa/US08_PUBCOMB.pap:*
- 9: /cgn2_6/prodata/1/pubaa/US09_PUBCOMB.pap:*
- 10: /cgn2_6/prodata/1/pubaa/US09B_PUBCOMB.pap:*
- 11: /cgn2_6/prodata/1/pubaa/US09C_PUBCOMB.pap:*
- 12: /cgn2_6/prodata/1/pubaa/US09_NEW_PUB.pap:*
- 13: /cgn2_6/prodata/1/pubaa/US10A_PUBCOMB.pap:*
- 14: /cgn2_6/prodata/1/pubaa/US10B_PUBCOMB.pap:*
- 15: /cgn2_6/prodata/1/pubaa/US10C_PUBCOMB.pap:*
- 16: /cgn2_6/prodata/1/pubaa/US10_NEW_PUB.pap:*
- 17: /cgn2_6/prodata/1/pubaa/US60_NEW_PUB.pap:*
- 18: /cgn2_6/prodata/1/pubaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1310	42.0	707	12	US-10-183-687-352
2	1296	41.6	710	12	US-10-183-687-356
3	1131	36.3	692	12	US-10-183-687-530
4	1119.5	35.9	553	12	US-10-183-687-362
5	1111.5	35.6	561	12	US-10-424-599-234179
6	1075.5	34.5	486	12	US-10-183-687-354
7	1047.5	33.6	574	12	US-10-225-066A-356
8	1047.5	33.6	574	12	US-10-225-066A-856
9	1047.5	33.6	574	12	US-10-225-066A-912
10	1047.5	33.6	574	12	US-10-374-780A-334
11	1040.5	33.4	516	12	US-10-183-687-428
12	1011	32.4	538	12	US-10-024-632-11
13	1010.5	32.4	538	12	US-10-183-687-487
14	1003	32.2	470	12	US-10-183-687-485
15	999	32.0	512	12	US-10-183-687-483

16	997.5	32.0	540	12	US-10-183-687-429
17	997	32.0	566	12	US-10-183-687-529
18	994	31.9	557	12	US-10-183-687-491
19	990.5	31.8	510	12	US-10-424-599-243772
20	986	31.6	510	12	US-10-183-687-348
21	977	31.3	558	12	US-10-225-066A-1014
22	977	31.3	558	12	US-10-374-780A-2496
23	974.5	31.2	498	12	US-10-225-066A-382
24	974.5	31.2	498	12	US-10-374-780A-346
25	971.5	31.1	639	12	US-10-437-963-151728
26	971	31.1	530	12	US-10-183-687-358
27	966	31.0	700	12	US-10-437-963-139609
28	964.5	30.9	469	12	US-10-183-687-338
29	961	30.8	492	12	US-10-183-687-489
30	960	30.8	557	12	US-10-424-599-213367
31	957.5	30.7	384	12	US-10-183-687-360
32	956.5	30.7	460	12	US-10-425-114-38886
33	951.5	30.5	642	12	US-10-024-632-6
34	949.5	30.4	484	12	US-10-183-687-334
35	948	30.4	410	12	US-10-425-114-40793
36	947.5	30.4	459	12	US-10-425-114-47823
37	947	30.4	264	12	US-10-425-114-53727
38	943.5	30.3	489	12	US-10-183-687-332
39	942	30.2	663	12	US-10-024-632-9
40	942	30.2	669	12	US-10-024-632-9
41	941	30.2	415	12	US-10-225-066A-592
42	941	30.2	415	12	US-10-374-780A-400
43	935	30.0	466	12	US-10-437-963-122829
44	934.5	30.0	354	12	US-10-425-114-36848
45	930	29.8	485	12	US-10-183-687-421

ALIGNMENTS

RESULT 1

US-10-183-687-352
; Sequence 352 Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Ramodu Orclayo O.
; APPLICANT: Harvell, Iselle T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 352
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Glycine max
US-10-183-687-352

Query Match 42.0%; Score 1310; DB 12; Length 707;
Best Local Similarity 45.6%; Pred. No. 3.4e-104;
Matches 352; Conservative 77; Mismatches 139; Indels 168; Gaps 29;
4 NWLGSLSPYEQNHKDVYSSTT-----TTVDVAGEYCYDPTAASDESSAIQ 52

```

; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 356
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Glycine max
US-10-183-687-356

Query Match 41.6%; Score 1296; DB 12; Length 710;
Best Local Similarity 45.5%; Pred. No. 5.5e-103; Indels 176; Gaps 29;
Matches 324; Conservative 66; Mismatches 146;

QY 4 NMLGFSLSPEY-----QNHR-----KDYISSTTTTVDVAGEYCYDPTAASDESSA 50
Db 5 NLLGFSLSPEHPSQDSHQSOTAPRSCFNPDPGISST-----DVAGD-CFDLTS---DSTP 55
QY 51 ICTSPSPGVVDAFTDRNNSH-SRDWDIN-----GCACNN---IHN---DEQDGP 95
Db 56 HLLNUPS-YG-IYEAHRNSNN.IHTDWRKYNNSQNLGLTSCSNQNNHNRHQOQQOOP 113
QY 96 KLENFLGRITTTTYNTINEN-VGDGSGGCV-----GGDGG-----GSLGLSM 137
Db 114 KLENFLGHS--FGEHEQPYGCSASTYWPAPQVPLAGGGGSGNSSTNSSSIGLSM 171
QY 138 IKTWLRNQPVDVNDQENGAAG-----LSLSMN-----SCDNNND 177
Db 172 IKTWLRNQPHENNNNNNNNESGNSRSVQOTLSLSNSTGSSQSTSLPLLTASVDNGES 231
QY 178 SNNN-----VVAQGTIDDSVEATPKTIESFGORTSYRGVTRHRWTGRYEHLWDNSC 232
Db 232 SDNQKPHTTAALDTTGTCAIETAPKSIDTFCORTSYRGVTRHRWTGRYEHLWDNSC 291
QY 233 KREGOTRGRQVYLGYYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKHMTR 292
Db 292 RREGOTRGRQVYLGYYDKEEKAARAYDLAALKYWGTTTTTTFPIHYEKELEEMKHMTR 351
QY 293 QYVASLRKSSGFRGASIVRGVTRHGHGRWQARIQVAGNKDLYLGTFTGTOEAAEA 352
Db 352 QYVASLRKSSGFRGASIVRGVTRHGHGRWQARIQVAGNKDLYLGTFTGTOEAAEA 411
QY 353 YDIAAKPRGLTAVTNFDMNRYNKAILESPLSPIGSAAKELKE-----ANR 399
Db 412 YDVAAKPRGLSAVTFDMRSYDVKSIESTTLPIGGAAKELKOMEQVELRVENVHRADQ 471
QY 400 PVPMMNNTSVSESENSASQWNAVQHOGVDLSLLHGHQERYNGYNGNLSSEA 459
Db 472 EDHSSIMSHLTQGINNYAAGGTATTATHHNHWNALAFHQPOPCCTTHYPYQORIN--- 527
QY 460 RACFKQEDDQ-----HHFLSNTQSL---MTNIDHQSVDSDSVT 495
Db 528 --WCKQEQNSDASHLSYSVDIHQIQLGNGGTHNFHTNSGLHPMLSMD---SASIDNSS 582
QY 496 VCGNVV--GYGYQGF-AAPV-NCDAYAASEFDYN----- 526
Db 583 SSNSVYDYGCGGGYVPMGTTTTTVVANDGDQPRSNHGFQDNEIKALGYESVYGSTT 642
QY 527 -----ARNHYFFAQOQQTQSPGDFPAAAMTNVGSNNYHGGEGGEVAPT 572
Db 643 DPHYHARNLVYLTOQPSS-----VDAVKASAYDQGSACNTWVPT 683

```

RESULT 3

```

US-10-183-687-530
; Sequence 530, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd

```

RESULT 2

```

US-10-183-687-356
; Sequence 356, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B1458 US NA
; CURRENT APPLICATION NUMBER: US/10/183.687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29

```

```
APPLICANT: Kinney, Tony
APPLICANT: Klein, Ted
APPLICANT: Li, Changjiang
APPLICANT: Oliveira, Igor Cunha
APPLICANT: Sakai, Hajime
APPLICANT: Shen, Bo
APPLICANT: Tarczynski, Mitchell C.
TITLE OF INVENTION: Alteration Of Oil Traits In Plants
FILE REFERENCE: BB1458 US NA
CURRENT APPLICATION NUMBER: US/10/183,687
CURRENT FILING DATE: 2002-06-27
PRIOR FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Microsoft Office 97
SEQ ID NO 530
LENGTH: 692
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: gi 2016:013
US-10-183-687-530

Query Match      36.3%; Score 1131; DB 12; Length 692;
Best Local Similarity 40.1%; Pred. No. 1e-88;
Matches 278; Conservative 65; Mismatches 125; Indels 226; Gaps 23;

Qy      3  NNWLGFSLSPYEONHHRKDVYSTTTTVDVAGEYCYDPTAADESASAIQTSPPSPFGV 62
Db      5  NNWLAFSLSPQDQLPP-----SQNSTLISAAA-----TTTACDSTGDCVFNIP----- 50
Qy      63  VDFTDRDNNHSHRDWDINGCACNNIHNDQDQPKLENFLGRITTTIYNTNENVDGQ----- 117
Db      51  -----QDWSRSGSLSALVAE-----PKLEDFLGGSISFSEQQHHGKGVIPS 94
Qy      118  SSGSCY-----GGG-----DG 128
Db      95  SAAACYASSGSSGYLYPPSSSSQLQFADSVWATSPVVAHDVGSGGGMVSAASAAAAS 154
Qy      129  GGSGLGLSMIKTWLRNOPVDNVDNQENGNAAGLSLSMNS----- 168
Db      155  GNGGIGLSMIKWMLRSQAP-----QPAQALSLSMNWAGTTTAQGGGAMALLAGAGE 206
Qy      169  -----STSC-----DNNDSNNNVAAQKTIDDS-----V 193
Db      207  RGRTPASESLSTSAHGATTATMAGGRKEINEEGSGSAGAVAVGSSGSGAVVEAGAA 266
Qy      194  EATPKTIESFGORTSIYRGVTRHRTGRYEHLWNSCKREGQTRKGRGVYLGVDKKE 253
Db      267  AAARKSVDFGQRTSIYRGVTRHRTGRYEHLWNSCRREGQTRKGRQ---GGYDKEE 323
Qy      254  KAARAYDLALKYGGTTTTNPFMSYEKEVEEMKMTROYVYASLRKSSGFSRGASII 313
Db      324  KAARAYDLALKYGGTTTTNPFVNNYEKELEEMKMTROEFVASLRKSSGFSRGASII 383
Qy      314  RGVTRHHQHGQWQARIQVAGNKDLYLGTGTOEEAAEAYDIAAKPRGLTAVTFDNR 373
Db      384  RGVTRHHQHGQWQARIQVAGNKDLYLGTSTQEEAAEAYDIAAKPRGLNATVTFDMSR 443
Qy      374  YNVKAILSPSLPGSAAKRLKEANRVPV-SMMISMNVS-----ESENSASGW 421
Db      444  YDVKSLDLSAALPVGTAARKLKDAEAAAAYDVGRVIAHSLGGDGAAYAAHYHHHSHSAAA 503
Qy      422  QNAAVQ-----HHQGVDSLHQQHOERYNGYNGNLSSESARACPKOEDDHHFLS 474
Db      504  PTIAFQAAAAAPPAAAG-----LYHPYAQPLRGW-----C-KOEQD-HAVIA 543
Qy      475  NTQSLMTNIDHQSSVSDSDSVTCGVNGVYGGYOGFAAPVNCDAVAASEFDYNARNHYIFA 534
Db      544  AAHSIQ-DLHH-----LNLGAAAAA-----HDFFS 567
```

```
Qy      535  QQQQTQSSPGDFFPAAMTNVGSN-MYYHGGGG 567
Db      568  QAMQQQHGLGSIIDNASLEHSTGNSNVVYNGDNGG 601
```

RESULT 4

```
US-10-183-687-362
; Sequence 362, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Catoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BB1458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 362
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Glycine max
US-10-183-687-362
```

```
Query Match      35.9%; Score 1119.5; DB 12; Length 553;
Best Local Similarity 43.5%; Pred. No. 7.2e-88;
Matches 275; Conservative 77; Mismatches 147; Indels 133; Gaps 23;
```

```
Qy      1  MNNWLGFLSPYEONHHRKDVYSTTTTVDVAGEYCYDPTAADESASAIQTSPPSPFG 60
Db      1  MNNWMLSPFLS-----PTHSLPAHDLQATQYHOPS 31
Qy      61  VVVDFTDRDNNHSHRDWD-INCCACNNIHNDQDQPKLENFLGRITTTIYNTNENVDGSG 119
Db      32  LGLVNNMDNPPQNHDMNLINTHSSNEI-----PKVADFLG-----VSKSENQSDLLA 79
Qy      120  SGCYGGGGGGGSLGLSMIKTWLRNOPVDNVDN--QENGNA-AKGLSLSMNS--STSCDN 174
Db      80  LNEIHSNDSLYLFTNNSLVP--MQNPVLDTPSNEYQENANSNLQSLTTLMSGKDKSTCET 137
Qy      175  NNDSDNNVVAQKTIDDSVEATPKTIESFGORTSIYRGVTRHRTGRYEHLWNSCKR 234
Db      138  SGENSTNT-----TVEVAPRRLTDFGORTSIYRGVTRHRTGRYEHLWNSCKR 188
Qy      235  EGQTRKGRGVYLGVDKKEKAARAYDLALKYGGTTTTNPFMSYEKEVEEMKMTROE 294
Db      189  EGSRKGRGVYLGVDKKEKAARAYDLALKYGGTTTTNPFISNYKELDEVMKMTROE 248
Qy      295  YVASLRKSSGFSRGASIIYRGVTRHHQHGQWQARIQVAGNKDLYLGTGTOEEAAEAYD 354
Db      249  FVAATRRKSSGFSRGASMYRGVTRHHQHGQWQARIQVAGNKDLYLGTGTFSTEEAAEAYD 308
Qy      355  IAAIKPRGLTAVTFDNRVNNYKAILSPSLPI-GSAAKRLKEA-----NRVPVSMMI- 407
Db      309  IAAIKPRGLNATVTFDMSRYDVKAILSNLPIGGAAKRLKEAQALESRRKEEVIAG 368
Qy      408  SNNVSESENSASGWQNA--VOHH-----QGVDSLHLH-QHOE---RYNGY 448
Db      369  SSSTFYGTSASSRRLHAYPLMQHHRHQPQPOPPLLTQLNHDISSSHFSHQDPLHHQGYI 428
```

QY 449 YNGNLSSESARACFK-QEDDOHH--FLSNTQSL---MTNIDHOSSV-----DDSVTV 497
 Db 429 OTQLQHQQSGASSYSFQNNAQFYNGYLDONHPALLQGMNMGSSSSSVLENNSNNN 488
 QY 498 GNVVGGYGGQGFAPVNCDA-----YAASEFDYNARNHYFAQQOCTQOSPGGD 546
 Db 489 NNVGGFVG-SGFGMASNATAGTCTVCTABELGLVKVDYDMPAGGYGGRSAADSMQTSNGG- 546
 QY 547 FPAAMTNVGSNMYHGGGGEVAPFTVWND 578
 Db 547 -----VFTWVND 553

RESULT 5

US-10-424-599-234179
 ; Sequence 234179, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 234179
 ; LENGTH: 561
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_53488C.1.pap
 US-10-424-599-234179

Query Match 35.6%; Score 1111.5; DB 12; Length 561;
 Best Local Similarity 42.0%; Pred. No. 3.6e-87;
 Matches 273; Conservative 76; Mismatches 140; Indels 161; Gaps 22;

QY 1 MNNWLGFLSPYEONHHRKDVYSSTTTTVDVAGEYCDPTAASDESSAIQTSPFPFG 60
 Db 1 MNNWLSPLS-----PHTSSLPAHDLQATQYHQS 31
 QY 61 VVDAFTDNNSHSRDWD-INGCANNIHNDQGPXLENPLGRTTTIYNTNENVDGSG 119
 Db 32 LGLVNNNDNPFQNHDMNLINTHSSNEI-----PKVADPLG-----VSKSENQSDLA 79
 QY 120 SGCVGGGSGGSLGLSMIKTWLRNPVDNVDN--QENGNA-AKGLSLSMNS--STSCDN 174
 Db 80 LNEHSDSDYLFNTNSLVP--MNPVLDTPSNYQENANSLQSLTSMGSGKDSCTET 137
 QY 175 NDSNNNVVAGCKTIDDSVEATPKKTESFGQRTSIYRGVTRHRWTGRYFAHLWDSCKR 234
 Db 138 SGENSTNT-----TTTTVEVAPRTLDTFGQRTSIYRGVTRHRWTGRYFAHLWDSCKR 192
 QY 235 EGOTRKGQVYLGVDKEEKAARYDLAALKYWGTTTTTTFPMSYEKEVEEMKHTROE 294
 Db 193 EGQGRKGQVYLGVDKEEKAARYDLAALKYWGTTTTTTFPMSYEKEVEEMKHTROE 252
 QY 295 YVALRRKSSGFSRGASIVRGVTRHHQGRWQARIGVAGNKDLYLGTFTQBEAAEAYD 354
 Db 253 FVAALRRKSSGFSRGASIVRGVTRHHQGRWQARIGVAGNKDLYLGTFTQBEAAEAYD 312
 QY 355 IAAIKFRGLTAVTFDMNRYNKAILESPLPI-GSAAKRLKEA-----NRPVPSMMMS 408
 Db 313 IAAIKFRGLNATVTFDMNRYNKAILESPLPIGCGGAKEKLAQALESSRKEEMIALG 372
 QY 409 NVSESENASGQWNA--VQHH-----QGVDLSL-----LHO-----439
 Db 373 SSTFQYGTSSNSRLHAYPLVQCHHHQFEQPOPLTLQNHDISSHFHQDPLHQGYIQ 432

QY 440 ---HQER-----YNGYYNGNLSSESARACFKQEDDOHHFLSNTQSLMTN 482
 Db 433 LQHQQSGSSSYSPQNNINNAOFYNGYNL-----QNHPPALQKXNNG 478
 QY 483 IDHQSVDSDSVTCNVGVYGGYQGFAPVNCDA-----YAASEFDYNARNHYFA 533
 Db 479 SSSSSSVLENNSNNNNTNNVGVFG-SGFGMASNATSGNTVGTAEELGLVKVDYDMPGGY- 536
 QY 534 AQOQCTQOSPGGDPAA-----MTNNVGSNMYHGGGGEVAPFTVWND 578
 Db 537 -----CGWSAAAAAESMQTSNSG-----VFTWVND 561

RESULT 6

US-10-183-687-354
 ; Sequence 354, Application US/10183687
 ; Publication No. US20030204870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Allen, William B.
 ; APPLICANT: Cahoon, Rebecca
 ; APPLICANT: Egelbaum, Sabine
 ; APPLICANT: Famodu, Omolayo O.
 ; APPLICANT: Harvell, Leslie T.
 ; APPLICANT: Jones, Todd
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Klein, Ted
 ; APPLICANT: Li, Changjiang
 ; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczynski, Mitchell C.
 ; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
 ; FILE REFERENCE: B01458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; PRIOR FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 354
 ; LENGTH: 486
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 US-10-183-687-354

Query Match 34.5%; Score 1075.5; DB 12; Length 486;
 Best Local Similarity 50.2%; Pred. No. 3.8e-84;
 Matches 240; Conservative 45; Mismatches 98; Indels 95; Gaps 12;

QY 167 NSTSCDNNNSNNNVAGCKTIDDSVEATPKKTESFGQRTSIYRGVTRHRWTGRYFAH 226
 Db 4 NGESSDNRKP--NTSALDSTQTGAETAPRKSIDTFQRTSIYRGVTRHRWTGRYFAH 61
 QY 227 LWDNCKEKGQTRKQVYLGVDKEEKAARYDLAALKYWGTTTTTTFPMSYEKEVEE 286
 Db 62 LWDNCKEKGQTRKQVYLGVDKEEKAARYDLAALKYWGTTTTTTFPMSYEKEVEE 121
 QY 287 MKHMTROEYVALSRKSSGFSRGASIVRGVTRHHQGRWQARIGVAGNKDLYLGTFTQ 346
 Db 122 MKHMTROEYVALSRKSSGFSRGASIVRGVTRHHQGRWQARIGVAGNKDLYLGTFTQ 181
 QY 347 EBAEAYDIAAKPRLTAVTFDMNRYNKAILESPLPIGSAKRLKE-----396
 Db 182 EBAEAYDIAAKPRLTAVTFDMNRYNKAILESPLPIGSAKRLKE-----396
 QY 397 ---ANRPVPSMMMSNNVSESENASGQWNAVQHHQGVDSLHQQHQRVNGYYNGN 453
 Db 242 GHRADQVDHSHIIMSSHLTQGINNVAGGTAT--HHNWHNAFAHQPOCTTMEHPYQOR 299
 QY 454 LSSESARACFKQEDDO-----HHFLSNTQSL--MTNIDHOSSV 489
 Db 300 IN-----WCKEQQDNDAPHSLSYSDIHQLOLGNNGTHNFFHTNSGLHFLMSWD---SA 351

QY 490 SDDSVTCGVNY--GYGGYQGF-AAFV-NCDAAYASEFDYN----- 526
Db 352 SIDNSSNSVYDYGCGGGYNWMPGTTTAVASDGDQONPRSNHGFCDNEIKALGYES 411
QY 527 -----ARHYFAQQOQOQOQSGGDFPRAATNNVGSNNMYHGGEGGVAPT 572
Db 412 VYGSATDSYHAHARNLYLTQOO-----SSSVDTVKASAYDOQASACNTWVPT 458

RESULT 7
US-10-225-066A-366
; Sequence 366, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 366
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-366

Query Match 33.6%; Score 1047.5; DB 12; Length 574;
Best Local Similarity 38.7%; Pred. No. 1.3e-81;
Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;

QY 2 NNNLGF-----SLSPYEONHHRKDYVSTTTTVDVAGEYCYDPTAASDESSAIQTS 54
Db 3 SNNLGFPLSPNNSLPPHEYN----- 24

QY 55 FPSFPGVVDAFTDNNKSHSRDWDINGACNNIHDEQDG----PKLENFLGRTTTTYNT 110
Db 25 ----LGLVSDFM--DNFQTOEWNM-----INPHGGGGDEGGEVPKVADFLGVSKPDENQ 73

QY 111 NENVGDSGGGCGYGGGCGGSLGSMIKTWLRNPQV-----DNVDN-----QENGNA 158
Db 74 SNHLVAYNDSVYFHTN-----SLMPSVQSNVDVVVAACDNTNPNSSYHELQESAHN 125

QY 159 AKGLSLNNSSTSCDNNDSNNVV---AQCKTIDDSVE-----ATPKKTIESFGQ 206
Db 126 LQSLTSLMGTT-----AGNVVDVKASPETTGDNAGGALAVVETATPRALDTFGQ 177

QY 207 RTSIYRGVTRHWTGRYEHLWDNSCKREGQTRKGRQVYLGGYDKEEKAAPAYDLAALKY 266
Db 178 RTSIYRGVTRHWTGRYEHLWDNSCRREGQSRGRQVYLGGYDKEDKAARSYDLAALKY 237

QY 267 WGTITTTNPPMSYEKEVEEMKMTROEYVVASLRKSSGFSRGASIVRGVTRHHQHGRWQ 326
Db 238 WGPSTTTNFTTYEKEVEEMKMTKQEFVAAAIRKSSGFSRGASMYRGVTRHHQHGRWQ 297
QY 327 ARIGRVAGNKDLYLGTFTGTEEAAYDIAAIAIKFRGLTAVTNFDMRYNVKALIESSLP 386
Db 298 ARIGRVAGNKDLYLGTFTGTEEAAYDIAAIAIKFRGLNAVTFNFEINRYDVKALIESSTLP 357
QY 387 I-GSAAKRLKEA-----NR 399
Db 358 IGGAAKRLKEAQALESRRKREAEIMIALGSSFYQGGSGSTSGSSTSRLLQLQPYPLSIQ 417

QY 400 PVPMSMISNNVSENSASGQNAVAQHGGVDLSLLHQHQRERYNGYNGNLSSESA 459
Db 418 PLEPFLSLQNNDISHYNNNAHDSSFNHHSYIQ-TQLHLHQITNNVYLQOQSSQSOQLY 476

QY 460 RACFKQEDDQHHFLSNTOSLMTNIDHQ--SSVSDSDSVTCGVNVGYGYQGFAPV-NCD 516
Db 477 NAYLHSPALLHGLVST-SIVDNNNNNGSGGSYNTAAFLGN---HGIGIGSSSTVGSTE 532

QY 517 AYAASEFDYNARNHYFAQQOQOQOQSGGDFPAAATNNVGSNNMYHGGEGGVAPTFTVW 576
Db 533 EFPVTKTDYD-----MPSDDGTGGYSGWTSSEVQGSNPGG-----VFTW 572

QY 577 ND 578
Db 573 NE 574

RESULT 8
US-10-225-066A-856
; Sequence 856, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 856
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-856

Query Match 33.6%; Score 1047.5; DB 12; Length 574;
Best Local Similarity 38.7%; Pred. No. 1.3e-81;
Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;

QY 2 NNNWLG-----SLSPYEQNHHRKDVYSSTTTTVDVAGEYCYDPTAASDESSAIOTS 54
 Db :|||||
 QY 3 SNNWLGFLPLSPNNSSLPPEYH----- 24
 Db :|||||
 QY 55 FSPFPGVVVDAFTRDNNSHRWDINGCACNNIHNDDEODG---PKLENFLGRITTIYNT 110
 Db :|||||
 QY 25 -----LGLVSDHM--DNPFQTQEWNN-----INPHGGGDEBGEVPEKVAADFLGVSKPDENQ 73
 Db :|||||
 QY 111 NENVGDSGGCGYGGGDSGLSLMTKWLNRNPV-----DNVDN-----QENGNA 158
 Db :|||||
 QY 74 SNHLVAYNDSDYFHTN-----SLMPSVQSNVDVVAACDSNTPNNSSVHELQESAHN 125
 Db :|||||
 QY 159 AKGLSLSMNSSTSCDNNDSNNVV-----AQGTIDDSVE-----ATPKKIESFGQ 206
 Db :|||||
 QY 126 LOSLILSMGT-----AGNVVDKASPSSETTGDNASGGALAVVETATPRRALDTPGQ 177
 Db :|||||
 QY 207 RTSIYRGVTRHRWTGRYEHLWDNSCKREGOTRKGQVYLGKYDKEKAAAYDLAALKY 266
 Db :|||||
 QY 178 RTSIYRGVTRHRWTGRYEHLWDNSCRREGOSRKGQVYLGKYDKEKAAARSYDLAALKY 237
 Db :|||||
 QY 267 MGTITTTNPFMSYKVEVEEMKMTROBYVASLRKSSGFSRGASIRYGVTRHRHQRWQ 326
 Db :|||||
 QY 238 WGPSTTTNPFITNYEKEVEEMKMTROBFVAAIRKSSGFSRGASMYRGVTRHRHQRWQ 297
 Db :|||||
 QY 327 ARIGRVAGNKDLYLGTFTGTEBAAEAYDIAAKFRGLTAVTNFDMNRYNVKAILSPSLP 386
 Db :|||||
 QY 298 ARIGRVAGNKDLYLGTFTSTEEBAAEAYDIAAKFRGLNAVTFNINRYDVKAILESSTLP 357
 Db :|||||
 QY 387 I-GSAAKPLKEA----- 399
 Db :|||||
 QY 358 IGGGAARLKEAQALESRRKREAEWIALGSSFYGGSGSTSGSTSRRLQLQPYPLSIQ 417
 Db :|||||
 QY 400 PVPMMWISNNVSESENSASQWAAVQHHCQVDSLHLOHOERYNGYNGNLSSEA 459
 Db :|||||
 QY 418 PLEPFLSQNDISHYNNNAHDSSFNHHSYIO-TQLHLHQNTNNYLQOQSSONSOLY 476
 Db :|||||
 QY 460 RACFKQEDDQHFLSNTQSLMTNIDHQ--SSVSDSVTVCGNVVGYGQYGFAPV-NCD 516
 Db :|||||
 QY 477 NAYLHSPALLHGLVST-SIVDNNNNGSGSGSYNTAAFLGN---HGIGIGSSSTVGSTE 532
 Db :|||||
 QY 517 AYAASEPDYNARNHYFAQQOQTOQSGGDFPAAMTNNVGSNNMYHGEGGEVAPFTVW 576
 Db :|||||
 QY 533 EFTVKTIDY-----MPSSDGTGYSWTSVQGSNPGG-----VFTMW 572
 Db :|||||

RESULT 9

US-10-225-066A-912
 ; Sequence 912, Application US/10225066A
 ; Publication No. US20030226173A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: RUECHMANN, Jose Luis
 ; APPLICANT: ADAM, Luc J
 ; APPLICANT: DUBELL, Arnold T
 ; APPLICANT: HEARD, Jacqueline E
 ; APPLICANT: PILGRIM, Marsha L
 ; APPLICANT: JIANG, Cai-Zhong
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: CREELMAN, Robert A
 ; APPLICANT: PINEDA, Omaira
 ; APPLICANT: YU, Guo-Liang
 ; APPLICANT: BROWN, Pierre E
 ; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: MB10036-2 US
 ; CURRENT APPLICATION NUMBER: US/10/225,066A
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 09/837,444

Query Match 33.6%; Score 1047.5; DB 12; Length 574;
 Best Local Similarity 38.7%; Pred. No. 1.3e-81;
 Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;
 QY 2 NNNWLG-----SLSPYEQNHHRKDVYSSTTTTVDVAGEYCYDPTAASDESSAIOTS 54
 Db :|||||
 QY 3 SNNWLGFLPLSPNNSSLPPEYH----- 24
 Db :|||||
 QY 55 FSPFPGVVVDAFTRDNNSHRWDINGCACNNIHNDDEODG---PKLENFLGRITTIYNT 110
 Db :|||||
 QY 25 -----LGLVSDHM--DNPFQTQEWNN-----INPHGGGDEBGEVPEKVAADFLGVSKPDENQ 73
 Db :|||||
 QY 111 NENVGDSGGCGYGGGDSGLSLMTKWLNRNPV-----DNVDN-----QENGNA 158
 Db :|||||
 QY 74 SNHLVAYNDSDYFHTN-----SLMPSVQSNVDVVAACDSNTPNNSSVHELQESAHN 125
 Db :|||||
 QY 159 AKGLSLSMNSSTSCDNNDSNNVV-----AQGTIDDSVE-----ATPKKIESFGQ 206
 Db :|||||
 QY 126 LOSLILSMGT-----AGNVVDKASPSSETTGDNASGGALAVVETATPRRALDTPGQ 177
 Db :|||||
 QY 207 RTSIYRGVTRHRWTGRYEHLWDNSCKREGOTRKGQVYLGKYDKEKAAAYDLAALKY 266
 Db :|||||
 QY 178 RTSIYRGVTRHRWTGRYEHLWDNSCRREGOSRKGQVYLGKYDKEKAAARSYDLAALKY 237
 Db :|||||
 QY 267 MGTITTTNPFMSYKVEVEEMKMTROBYVASLRKSSGFSRGASIRYGVTRHRHQRWQ 326
 Db :|||||
 QY 238 WGPSTTTNPFITNYEKEVEEMKMTROBFVAAIRKSSGFSRGASMYRGVTRHRHQRWQ 297
 Db :|||||
 QY 327 ARIGRVAGNKDLYLGTFTGTEBAAEAYDIAAKFRGLTAVTNFDMNRYNVKAILSPSLP 386
 Db :|||||
 QY 298 ARIGRVAGNKDLYLGTFTSTEEBAAEAYDIAAKFRGLNAVTFNINRYDVKAILESSTLP 357
 Db :|||||
 QY 387 I-GSAAKPLKEA----- 399
 Db :|||||
 QY 358 IGGGAARLKEAQALESRRKREAEWIALGSSFYGGSGSTSGSTSRRLQLQPYPLSIQ 417
 Db :|||||
 QY 400 PVPMMWISNNVSESENSASQWAAVQHHCQVDSLHLOHOERYNGYNGNLSSEA 459
 Db :|||||
 QY 418 PLEPFLSQNDISHYNNNAHDSSFNHHSYIO-TQLHLHQNTNNYLQOQSSONSOLY 476
 Db :|||||
 QY 460 RACFKQEDDQHFLSNTQSLMTNIDHQ--SSVSDSVTVCGNVVGYGQYGFAPV-NCD 516
 Db :|||||
 QY 477 NAYLHSPALLHGLVST-SIVDNNNNGSGSGSYNTAAFLGN---HGIGIGSSSTVGSTE 532
 Db :|||||
 QY 517 AYAASEPDYNARNHYFAQQOQTOQSGGDFPAAMTNNVGSNNMYHGEGGEVAPFTVW 576
 Db :|||||
 QY 533 EFTVKTIDY-----MPSSDGTGYSWTSVQGSNPGG-----VFTMW 572
 Db :|||||
 QY 577 ND 578
 Db :
 QY 573 NE 574

RESULT 10

US-10-374-780A-334
 ; Sequence 334, Application US/10374780A

Publication No. US20040019927A1
GENERAL INFORMATION:
APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddle, James
APPLICANT: Brown, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omaira
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: WBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
PRIOR FILING DATE: 2003-02-25
CURRENT FILING DATE: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: Patent in version 3.2
SEQ ID NO 334
LENGTH: 574
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: GI793 (conserved domain in AA coordinates:179-255, 281-349)
US-10-374-780A-334

Query Match 33.6%; Score 1047.5; DB 15; Length 574;
Best Local Similarity 38.7%; Pred. No. 1.3e-81;
Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;
Qy 2 NNNWLGPF-----SLSPYEQNHKRDVYSTTTTVDVAGEYCYDPTAASDESSAQTQS 54
Db 3 SNNWLGFLSPNNSLPPEYHN-----24
Qy 55 FPSPFVGVVDAFTDRNNSHSDWDINGCACNHNHDEQDG-----PKLENFLGRITTTTYNT 110
Db 25 ----LGLVSDHM--DNPQTQEWNM-----INPHGGGDEGGEVVKVADFLGVSKPDENQ 73
Qy 111 NENVGDSGGCGYGGGCGGGLSMKTKWLRNQPV-----DNDVN-----QENGNA 158
Db 74 SNHLVAYNDSYFYHTN-----SLMPVQSNVDVVAAACDSNTPNNSVHELQESAHN 125
Qy 159 AKGLSLNNSSTSCDNNNDNNNVV---AOGKTIIDDSVE-----ATPKTIESFGQ 206
Db 126 LQSLTSLMGIT-----AGNVVDVKASPEITGDNAGGALAVVETATPRALDIFGQ 177
Qy 207 RTSIYRGVTRHRTWTRYEAHLWDNSCKREGQTRKGRQVYLGVDYKEKKAAPAYDLAALKY 266
Db 178 RTSIYRGVTRHRTWTRYEAHLWDNSCKREGQTRKGRQVYLGVDYKEDKAARSYDLAALKY 237

Qy 267 WGTITTTNFMSEYEKEVEEMKHMTRQEVYASLRRKSSGSRGASLYRGVTRHHQGRWQ 326
Db 238 WGPSTTTNFPITNYEKEVEEMKHMTRQEFVAAIRKSSGSRGASMYRGVTRHHQGRWQ 297
Qy 327 ARIGRVAGNKDLYLGTCTGTOEAAEAAYDIAAIAKFERGLTAVTNFDMNRYNKAILESPSLP 386
Db 238 ARIGRVAGNKDLYLGTSTEBEAAEAAYDIAAIAKFERGLNAVTFNFEINRYDVKAILESTLP 357
Qy 387 I-GSAAKPLKEA-----NR 399
Db 358 ICGGAATKLEAQALESRRKEAEMIALGSSFYGGGSGTSGSTSRLOQPYPLSIQQ 417
Qy 400 PVPSMMISMNNVSESENSASGQNAVQHQHGVLSLLHQHQERYNGYYNGNLSBSA 459
Db 418 PLEPFLSLQNDISHYNNNNNAHDSSFNHHSYIQ-TQLHLHQQTNNYLQQOSSQNSQQLY 476
Qy 460 RACFKQEDDQHHFLSNTQSLMTNIDHQ--SSVSDSDSVTCGVNVGVGYQGFAPV-NCD 516
Db 477 NAYLHNPALLHGLVST-SIVDNNNNNGSSGSYNTAAFLGN---HGIGIGSSSTVSGSTE 532
Qy 517 AYAASEFDYNARNHYVFAQQOQTQOSRGGDFPAAATNNVGSNNVYVHGGGGEVAPTFTVW 576
Db 533 EFTVTKTIDY-----MPSSDGTGGYSGWTSESVOGSPGG-----VFTMW 572
Qy 577 ND 578
Db 573 NE 574
RESULT 11
US-10-183-687-428
Sequence 428, Application US/10183687
Publication No. US20030204870A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Allen, William B.
APPLICANT: Cahoon, Rebecca
APPLICANT: Epeibaum, Sabine
APPLICANT: Famodu, Omolayo O.
APPLICANT: Harvell, Leslie T.
APPLICANT: Jones, Todd
APPLICANT: Kinney, Tony
APPLICANT: Klein, Ted
APPLICANT: Li, Changjiang
APPLICANT: Oliveira, Igor Cunha
APPLICANT: Sakai, Hajime
APPLICANT: Shen, Bo
APPLICANT: Tarczynski, Mitchell C.
TITLE OF INVENTION: Alteration Of Oil Traits In Plants
FILE REFERENCE: BB1458 US NA
CURRENT APPLICATION NUMBER: US/10/183,687
CURRENT FILING DATE: 2002-06-27
PRIOR APPLICATION NUMBER: 60/301,913
PRIOR FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Microsoft Office 97
SEQ ID NO 428
LENGTH: 516
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: gi 4836931
US-10-183-687-428
Query Match 33.4%; Score 1040.5; DB 12; Length 516;
Best Local Similarity 44.6%; Pred. No. 4.5e-81;
Matches 254; Conservative 70; Mismatches 114; Indels 131; Gaps 22;
Qy 86 NIHNSQDG---PKLENFLGRITTT-IYNTNEN-----VGDGSGSGCYGGGCGGSGSLGI 135
Db 3 NPHGGGGEVGPVKVADFLGVSKSGDHTDNLVYPYNDIHOQNASDYFQTN-----SL-L 57

QY 136 SMKTLRLNQPVDVNDQNGNAKGLSLSNSS-----TSCNNNDNS 179
Db 58 PTWVCASNAP--NNVLEQSAHNLQSLTSLMGSTGAAAAEAVTKASPAETSADNSSTT 116
QY 180 NNVAAGKTIIDSVATPKTTTSTFGORTSIYRGVTRHWTGRVYEAHLWDNSCKEKGOTR 239
Db 117 N--TSGGAI---VEATPRTLETFGORTSIYRGVTRHWTGRVYEAHLWDNSCKEKGOTR 170
QY 240 KGRQVYLGYYDKEEKAARAYDAAALKYMGTTTTTTFPMSEYEKEVEEMKMTROEYVASL 299
Db 171 KGRQ---GGYDKEEKAARAYDAAALKYMGTTTTTTFPMSEYEKEVEEMKMTROEYVASI 227
QY 300 RKSSGFSRGSIIYRGVTRHWHOGHQAIRIGVAGNKDLYLGTGTQTEAAEAYDIAIK 359
Db 228 RKSSGFSRGSIIYRGVTRHWHOGHQAIRIGVAGNKDLYLGTGTSTEEBAEAYDIAIK 287
QY 360 FGLTAVTNFDMRYNVKXAILESPLPI--GSAKELKEA-----NRPVPMMSNNVSE 413
Db 288 FGLTAVTNFDMRYNVKXAILESPLPI--GSAKELKEA-----NRPVPMMSNNVSE 413
QY 414 SENSASGQNAAR-----VQHOGV-----DLS----- 435
Db 348 -YGAASGSSVASSRLOLPYFLSIQPFELHHPQLTLQNNNDISQVHDSFSYIOT 406
QY 436 LLHQERYNGYNGGNSLSESARACFKQEDDQHHFLSNTQSLMTNIDHSSVDDSVT 495
Db 407 QLHLHQOQTNNYLQSSHTSQL-----YNAYLQSNPGLL-----HGFVSDNN-- 448
QY 496 VCGNVVGGYOGFPAAPVNCDAAYASEFDYNARHNYFAOQOQTCQSPGGFPAAMTNV 555
Db 449 ---NTSGFLGNGIGIGSSSTVGSSABEEFPVKVDY-----DMPPS----- 487
QY 556 GSNMTYHGGEGGAVP-----TFTVND 578
Db 488 GGATCYGWNSSGESAQSGNPGGVFTWNE 516
RESULT 12
US-10-024-632-11
; Sequence 11, Application US/10024632
; Publication No. US20020170093A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: He, Steve S.
; APPLICANT: Dotson, Stanton B.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION A
; FILE OF INVENTION: GROWTH AND USES THEREOF
; FILE REFERENCE: 38-21(51837)B
; CURRENT APPLICATION NUMBER: US/10/024,632
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,896
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-10-024-632-11
Query Match 32.4%; Score 1011; DB 13; Length 585;
Best Local Similarity 38.5%; Pred. No. 1.9e-78;
Matches 258; Conservative 70; Mismatches 162; Indels 180; Gaps 22;
QY 3 NNWLGFLSLP---VEQNHKXDVSTTTTVDVAGYCYVDPTAASDESAIOTSPSPF 59
Db 2 SNWLGFLSLPDLRIEDSGFGRDHGFFSPVPLRSDSLCL----- 40
QY 60 GWVVDATRONNSHRWDIN---GCACNINHDEQDGPKEFLG-----RTTIV- 108
Db 41 --VVDPRRSIADEWRVENGIGSATN-----EQGPKLEDFLCYNSPQETKAYC 93

QY 109 -----NT-----NENVGDGSGSCYCGGD----- 127
Db 94 GTHENQNTVPSPTRINNVNAPNYSS-----SGDAEAAAFNTPSSFTQTYRNNENPQTL 149
QY 128 GGGGSL-----GLSMIKTMLRNQPDV-----VDNQENGN-NA 159
Db 150 AGHSLQQDDPNHNRQSGVHHVPESATSVGFKSWLRQTPFGGKASGNETNNNFM 209
QY 160 KGLSLMNSSTSCDNNNDNNN--VVAQKTTIDD-----SVEATPKTIESFGQT 208
Db 210 QALSMTSP-----SRNGFPALAPLEVVDNRKRPVGNLTRESVPRKSIDTFGQT 261
QY 209 SIYRGVTRHWTGRVYEAHLWDNSCKEKGOTRQVYLGYYDKEEKAARAYDAAALKYMG 268
Db 262 SQYRGVTRHWTGRVYEAHLWDNSCKEKGOTRQVYLGYYDKEEKAARAYDAAALKYMG 321
QY 269 TTTTTFPMSEYEKEVEEMKMTROBYVASLRSKSGFSRGSIIYRGVTRHWHOGHQAIR 328
Db 322 PTHINFPPLSTYEKELEEMKMTROEFVAHLRKSFGFSRGSIIYRGVTRHWHOGHQAIR 381
QY 329 IGRVAGNKDLYLGTGTQTEAAEAYDIAIKRPGLTAVTNFDMRYNVKXAILESPLPIG 388
Db 382 IGRVAGNKDLYLGTGTQTEAAEAYDIAIKRPGLTAVTNFDMRYNVKXAILESPLPIG 441
QY 389 SAAKRLKEANRPVPMMSNNVSESENSASGQNAARVQHOGVDSLHLHQERYNGY 448
Db 442 ELAKR-----SPKDTASIAPEYDYNCSASSAQPLLAIP----- 475
QY 449 YNGNLSSESARACFKQEDDQHHFLSNTQSLMTNIDHSSVDDSVTVCNVVGGYCG 508
Db 476 --SGEASDELADVMWTANSDEQQ-----CHQSTNTNDASLANSSRNSNPQSPKSGIG 528
QY 509 FAAPVNCDAAYASEFDYNARHNYFAOQOQTCQSPGGFPAAMTNVGSNNVYHGGEG 568
Db 529 LAS-----DRFGIGG-DYS--HHGYFSLGSKYEDGNSSTDNENRNLGNLGLVH----- 575
QY 569 VAPTFTVND 578
Db 576 KIPMFALMNE 585
RESULT 13
US-10-183-687-487
; Sequence 487, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 487
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-10-183-687-487

```
Query Match. 32.4%; Score 1010.5; DB 12; Length 538;
Best Local Similarity 43.9%; Pred. No. 1.9e-78;
Matches 249; Conservative 51; Mismatches 138; Indels 129; Gaps 18;

QY 3 NNWLGLSLP--YEQHHRKQVYSSTTTVDVAGYCYVDPTAASDESSAIOSTSPSPG 60
DB 7 NNSLAFSLSHHHQNH--LSFFSSSNPTLFOA-----FNTSAA-----AAVPTTPPA-- 54
QY 61 VVDAFTRDNNSHSRDWDINGCACNNIHNDQD-----GPKLENFLGRITTIYNTNENV 114
DB 55 -----NLSAEKADAGGEPATTASATTDLMSFKGPKLEDFLGSSIT----- 97
QY 115 GDGSGCYGGGGGSLGLSMIKTWRNQPVDVNDQENGNAAGLS-----LSMNSSTS 171
DB 98 -----PGVGAPQLQLHHHNFSTQPTVSDNDIYDSELKTIAAIFLRGFSSEQ 147
QY 172 CDNNDSNNVAAQKTTIDSVBATEPKKTIIESFGQRTSIYRGVTRHRWTCRYEAHLWDNS 231
DB 148 IDHKQOQOAPIA-----EPAPKESVDITFGQRTSIYRGVTRHRWTCRYEAHLWDNS 198
QY 232 CKREGQTRKRGQVYLGKYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMT 291
DB 199 CRREGOSRKRQVYLGKYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMT 258
QY 292 ROEVASLRKSGFSRGASIRGVTRHHOHWQARIGRVAGNKDLYLGTFTQOEAAE 351
DB 259 ROEVASLRKSGFSRGASIRGVTRHHOHWQARIGRVAGNKDLYLGTFTQOEAAE 318
QY 352 AYDIAIKPGLTAVTNFDMNRYNKAILESPLSGAARLKEANRPVSMNNTNNV 411
DB 319 AYDIAIKPGLTAVTNFDMNRYNKAILESPLSGAARLKEANRPVSMNNTNNV 371
QY 412 SESENSASG-----WQNAAVOHQGVDSL-----LH 438
DB 372 TDEERHSGSGMTFTSQPAASTVGFPAINKOPSDYWSNVLGYNDTGASLSAKNPF 431
QY 439 QHORYNGYVY-----NGGN--LSSESARAFK-----QEDDOHFLSNQSLMT-- 481
DB 432 HSPNGSGFQSPSGFCTGCTGNGSNGSNGSNGSNGSNGSNGSNGSNGSNGSNGSNGS 491
QY 482 ---NIDHQSVDSDSVTVGNVGYGG 505
DB 492 VLANSNDYESSSS-----GYGG 508

RESULT 14
US-10-183-687-485
; Sequence 485, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
```

```
; SEQ ID NO 485
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Catalpa sp.
US-10-183-687-485

Query Match. 32.2%; Score 1003; DB 12; Length 470;
Best Local Similarity 45.7%; Pred. No. 6.8e-78;
Matches 235; Conservative 49; Mismatches 94; Indels 136; Gaps 16;

QY 2 NNNMLGFLSLPYEQHHRKQVYSSTTTVDVAGYCYVDPTAASDESSAIOSTSPSPG 61
DB 7 HQNWLAFSLS-----NHS-----LFHPSFNPNSSSC-----GV 35
QY 62 VVDAFTRDNNSHSRDWDINGCACNNIHND--EQDGPKELENFLGRITTIYNTNENV 119
DB 36 V-----NOENG-----AGGAADHHLVSLASGPKLEDFLGASC-----GSS 73
QY 120 SCYGGGGGGGGGSLGLSMIKTWRNQPVDVNDQENGNAAGLSLMSNSTSCDNNDSN 179
DB 74 GGHITQYDGAQGVNISDVEIY-----GDSELKTIAAFLRGLSHERSDDSQ 120
QY 180 NNVAQKTTIDSVBATEPKKTIIESFGQRTSIYRGVTRHRWTCRYEAHLWDNSCKREGQTR 239
DB 121 KQTAVAAP-----EPAPKAAETFGQRTSIYRGVTRHRWTCRYEAHLWDNSCKREGQTR 175
QY 240 KGRQVYLGKYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMTROEYV 299
DB 176 KGRQVYLGKYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMTROEYV 235
QY 300 RKSGFSRGASIRGVTRHHOHWQARIGRVAGNKDLYLGTFTQOEAAEAYDIAIK 359
DB 236 RKSGFSRGASIRGVTRHHOHWQARIGRVAGNKDLYLGTFTQOEAAEAYDIAIK 295
QY 360 FGLTAVTNFDMNRYNKAILESPLSGAARLKEANRPV-----SMMNISNNV 411
DB 296 FGLTAVTNFDMNRYNKAILESPLSGAARLKEANRPV-----SMMNISNNV 354
QY 412 S-ESENSAS-----GQNAAVOH-----HQGVDSLH----- 438
DB 355 SFPSPNSAAAINFALPIKQDHSLSLWALGYQNSITNTNTPNTSVSLFHDNSFMBFQ 414
QY 439 -----QHORYNGYVYNGNLSSES 458
DB 415 ANNEGFFNGGYFQHQE-----NGTNTSTS 441

RESULT 15
US-10-183-687-483
; Sequence 483, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
```

